

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 897.829 Seconds
(without alignments)
11183.052 Million cell updates/sec

Title: US-09-513-888c-1_COPY_112_456

Perfect score: 345

Sequence: 1 tcgcagtagtaagctgcgcaaa.....gcaggagaaagcaaaatggg 345

Scoring table: IDENTITY_NUC

Gapop 10_0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_hgt_hum.*

31: em_hgt_inv.*

32: em_hgt_other.*

33: em_hgt_mus.*

34: em_hgt_pln.*

35: em_hgt_rtd.*

36: em_hgt_mam.*

37: em_hgt_vrt.*

38: em_sy.*

39: em_hgt_hum.*

40: em_hgt_mus.*

41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	345	100.0	9108	9	AF123653	AF123653 Homo sapi
2	345	100.0	173264	2	AC025853	AC025853 Homo sapi
3	286	82.9	633	9	AF123654	AF123654 Homo sapi
4	286	82.9	1515	9	AF123656	AF123656 Homo sapi
5	286	82.9	1614	9	AF123655	AF123655 Homo sapi
6	286	82.9	1692	9	AF123657	AF123657 Homo sapi
7	286	82.9	1722	9	AF123658	AF123658 Homo sapi
8	286	82.9	5492	9	AF123659	AF123659 Homo sapi
9	232.8	67.5	227884	2	AC114995	AC114995 Mus muscu
10	232.8	67.5	263546	2	AC099416	AC099416 Mus muscu
11	230.8	66.9	191210	2	AC108987	AC108987 Rattus no
12	106	30.7	231	9	AF123652	AF123652 Homo sapi
13	42.4	12.3	90348	1	AF497482	AF497482 Micromono
14	39.6	11.5	1981	6	AR103043	AR103043 Sequence
15	39.6	11.5	1981	6	AR103044	AR103044 Sequence
16	39.6	11.5	1981	6	AR103045	AR103045 Sequence
17	39.6	11.5	1981	6	AR103046	AR103046 Sequence
18	39.6	11.5	1981	6	AR103047	AR103047 Sequence
19	39.6	11.5	1981	6	AR103048	AR103048 Sequence
20	39.6	11.5	1981	6	E24180	E24180 Alpha-amyla
21	39.6	11.5	1981	6	E24181	E24181 Alpha-amyla
22	39.6	11.5	1981	6	E24182	E24182 Alpha-amyla
23	39.6	11.5	1981	6	E24183	E24183 Alpha-amyla
24	39.6	11.5	1981	6	E24184	E24184 Alpha-amyla
25	39.6	11.5	1981	6	E24185	E24185 Alpha-amyla
26	39.6	11.5	3496	1	PSEMFA	D10769 Pseudomonas
27	39.6	11.5	3496	6	E03105	E03105 DNA encodin
28	39.6	11.5	177559	4	AC091401	AC091401 Sus scrof
29	38.6	11.2	324	6	E15909	E15909 gDNA encodi
30	38.6	11.2	324	6	E30020	E30020 Mutated cyt
31	38.6	11.2	873	1	DVUPCC3	D31702 Sulfate-red
32	38.6	11.2	873	6	E37765	E37765 Process for
33	38.6	11.2	873	6	E37766	E37766 Process for
34	38.6	11.2	873	6	E64601	E64601 Process for
35	38.4	11.1	6058	1	AF279141	AF279141 Rhodococc
36	38.2	11.1	10295	1	AE012032	AE012032 Xanthomon
37	37.4	10.8	196050	1	AL646058	AL646058 Ralstonia
38	37.4	10.8	230138	14	AF232689	AF232689 Rat cytom
39	37	10.7	11044	1	AE012279	AE012279 Xanthomon
40	36.8	10.7	25970	1	SC2H4	AL031514 Streptomy
41	36.8	10.7	146884	2	AP004675	AP004675 Oryza sat
42	36.6	10.6	1884	10	AF050418	AF050418 Mus muscu
43	36.6	10.6	12795	1	AE004467	AE004467 Pseudomon
44	36.4	10.6	2267	6	AR195560	AR195560 Sequence
45	36.4	10.6	4800	6	AR106490	AR106490 Sequence

ALIGNMENTS

RESULT 1
AF123653

LOCUS

DEFINITION

ACCESSION AF123653

VERSION AF123653.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 9108)

Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,

Mori,M., Fidanza,V., Alder,H. and Croce,C.M.

The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

AF123653 Homo sapiens FEZ1 (FEZ1) gene, complete cds. 9108 bp DNA linear PRI 07-APR-1999

GI:4572463

----- Project Information

Center project name: L7454

Center clone name: 353_K_12

- * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

- * 1 39461: contig of 39461 bp in length
 * 39462 39561: gap of 100 bp
 * 39562- 108347: contig of 68786 bp in length
 * 108348 108447: gap of 100 bp
 * 108448 173264: contig of 64817 bp in length.

FEATURES

Location/Qualifiers

source

1..173264

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-353K12"

/clone_lib="RPC1-11 Human Male BAC"

BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others

ORIGIN

```

Query Match      100.0%; Score 345; DB 2; Length 173264;
Best Local Similarity 100.0%; Pred. No. 6.9e-78;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
Db 22393 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 22452

QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120
Db 22453 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 22512

QY 121 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 180
Db 22513 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 22572

QY 181 GATTACACGGCACTCTCCAGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCGG 240
Db 22573 GATTACACGGCACTCTCCAGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCGG 22632

QY 241 TCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGGTAAGCGGGGTGCG 300
Db 22633 TCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGGTAAGCGGGGTGCG 22692

QY 301 TCGCAGGCTAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 345
Db 22693 TCGCAGGCTAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 22737

```

RESULT 3

AF123654

LOCUS

AF123654 Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
 DEFINITION complete cds.

ACCESSION

AF123654

VERSION

AF123654.1

KEYWORDS

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 633)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,

Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

and its expression is altered in multiple human tumors
 Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..633

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="E264162"

1..633

/gene="FEZ1"

/note="alternatively spliced"

/codon_start=1

/product="FEZ1"

/protein_id="AAD23835.1"

/db_xref="GI:4572466"

/translation="MGVSLSLISGHSPHSHKCRASQYKLRKSSHLKLNRYSDGLLRF

GFSODSGHKSXKMGKSEDFYIKVQKARGSHDPDYVLTALSSDGLGGAGVDFDPST

PPKLMPPSNQIEMSGEKAIVATFKVILPRSGAILHSSPESASHOLHPAPPDPKPKQ

ELKPGLCGSAUSDRSNMSSLPTHSAGEPUEVDLEGADIPYEDIIATEI"

BASE COUNT 136 a 217 c 175 g 105 t

ORIGIN

```

Query Match      82.9%; Score 286; DB 9; Length 633;
Best Local Similarity 100.0%; Pred. No. 9.1e-63;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120
Db 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 180

QY 121 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 180
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGGCAAGGCGGGCTCCCATCACCA 240

QY 181 GATTACACGGCACTCTCCAGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCGG 240
Db 241 GATTACACGGCACTCTCCAGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCGG 300

QY 241 TCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 286
Db 301 TCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 346

```

RESULT 4

AF123656

LOCUS

AF123656 Homo sapiens clone D14 FEZ1 (FEZ1) mRNA, alternatively spliced,
 DEFINITION complete cds.

ACCESSION

AF123656

VERSION

AF123656.1

KEYWORDS

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1515)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,

Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
 MEDLINE 99199287
 PUBMED 10097140
 REFERENCE 2 (bases 1 to 1515)
 AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.
 TITLE Direct Submission
 JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA

FEATURES

Source

1..1515

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="D14"

1..1515

/gene="FEZ1"

1..1515

/note="alternatively spliced"

/codon_start=1

/product="FEZ1"

/protein_id="AAD23837.1"

/db_xref="GI:4572470"

/translation="MGVSLSLISGHFSHKRQAYKLRKSHLKKLNRYSDGLLRF
 GFSQSGHGKSSKMGKEDFFYIKVQKARGSHHPDYATLSSGLGGQAGVDFPST
 PKLMFNSQLMGSEKGAVRTAFKPLVPRSGAILHSSPESASHQHPADPKPEQ
 ELKPGICSLSDSGNSMSLPTHTSTSSYOLDPLVTPVGTSPRFGSAHNITOGIV
 LQSNMMSLKALSFDSGSKLGHNSKADKPCVSPSTDECSIQLEOKLLEREGA
 LQKQRFEEKELASLAYERPRRCRDEGPEPKGNKLAQKQKSORAQOVLHLQ
 VLQOKERQLESLMEQDLLETKLRSYERETSGPALETQWECVCKSGEIS
 LLKQOKESQTEVNAKSEILGLKQDKTRKLEGLRLTQDLGALRTKGLELEVC
 ENELQOSYVYVMTQNRORLEKALQQLARGDSAGEPLEVDLEGADIPYEDIATETI"

BASE COUNT 336 a 483 c 482 g 214 t

ORIGIN

Query Match 82.9%; Score 286; DB 9; Length 1515;

Best Local Similarity 100.0%; Pred. No. 9e-63;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60

DB 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCCAAAATGGGC 120

DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCCAAAATGGGC 180

QY 121 AAGAGCGAAGACTTCTTACATCAAGGTCAAGGTCAAGAAAGCCGGGGCTCCCATACCCA 180

DB 181 AAGAGCGAAGACTTCTTACATCAAGGTCAAGGTCAAGAAAGCCGGGGCTCCCATACCCA 240

QY 181 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240

DB 241 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286

DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346

RESULT 5

AF123655

LOCUS

DEFINITION

Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, alternatively

spliced, complete cds.

ACCESSION

AF123655

VERSION

AF123655.1

KEYWORDS

Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1614)

AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

MEDLINE 99199287

PUBMED 10097140

REFERENCE 2 (bases 1 to 1614)

AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

TITLE Direct Submission

JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA

FEATURES

Location/Qualifiers

1..1614

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="T8D145M4"

1..1614

/gene="FEZ1"

1..1614

/note="alternatively spliced"

/codon_start=1

/product="FEZ1"

/protein_id="AAD23836.1"

/db_xref="GI:4572468"

/translation="MGVSLSLISGHFSHKRQAYKLRKSHLKKLNRYSDGLLRF
 GFSQSGHGKSSKMGKEDFFYIKVQKARGSHHPDYATLSSGLGGQAGVDFPST
 PKLMFNSQLMGSEKGAVRTAFKPLVPRSGAILHSSPESASHQHPADPKPEQ
 ELKPGICSLSDSGNSMSLPTHTSTSSYOLDPLVTPVGTSPRFGSAHNITOGIV
 LQSNMMSLKALSFDSGSKLGHNSKADKPCVSPSTDECSIQLEOKLLEREGA
 LQKQRFEEKELASLAYERPRRCRDEGPEPKGNKLAQKQKSORAQOVLHLQ
 VLQOKERQLESLMEQDLLETKLRSYERETSGPALETQWECVCKSGEIS
 LLKQOKESQTEVNAKSEILGLKQDKTRKLEGLRLTQDLGALRTKGLELEVC
 ENELQOSYVYVMTQNRORLEKALQQLARGDSAGEPLEVDLEGADIPYEDIATETI"

BASE COUNT 367 a 500 c 522 g 225 t

ORIGIN

Query Match 82.9%; Score 286; DB 9; Length 1614;

Best Local Similarity 100.0%; Pred. No. 9e-63;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60

DB 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCCAAAATGGGC 120

DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCCAAAATGGGC 180

QY 121 AAGAGCGAAGACTTCTTACATCAAGGTCAAGGTCAAGAAAGCCGGGGCTCCCATACCCA 180

DB 181 AAGAGCGAAGACTTCTTACATCAAGGTCAAGGTCAAGAAAGCCGGGGCTCCCATACCCA 240

QY 181 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240

DB 241 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286

DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346

RESULT 6

AF123657

Qy	241	TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG		286
Db	301	TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG		346
RESULT 7				
LOCUS	AF123658	1722 bp	mRNA	linear PRI 07-APR-1999
DEFINITION	Homo sapiens clone G3612 FEZ1 (FEZ1) mRNA, alternatively spliced, complete cds.			
ACCESSION	AF123658			
VERSION	AF123658.1	GI:4572473		
KEYWORDS	Homo sapiens.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1	(bases 1 to 1722)		
AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.			
TITLE	The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)			
MEDLINE	99199287			
PUBMED	10097140			
REFERENCE	2	(bases 1 to 1722)		
AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA			
FEATURES	Location/Qualifiers			
source	1..1722			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="8"			
	/map="8p22"			
	/clone="G3612"			
gene	1..1722			
	/gene="FEZ1"			
CDS	1..1722			
	/gene="FEZ1"			
	/note="alternatively spliced"			
	/codon_start=1			
	/product="FEZ1"			
	/protein_id="AAD23839.1"			
	/db_xref="GI:4572474"			
	/translation="MGSVSLISGHSHFHSKRCASQYKLRKSHLKKLNRYSDGLLRF			
	GFSDQHQKSSKMGSGDFDFVIKVSQKARGSHDPDYALSSGDLGGQAGVDFDPST			
	PKLMFSPNQLMWSFKGAVRTPAKPVLPSGAILHSSPESASHOLHPAPDPKPEQ			
	ELKPLGCSALSDSGRNSMSLPTHSTSSYQLDPLVTPVGTSPRFGSGAHNITQGI			
	LQDSNMSLKALSFDSGGKLGHSNKADKPGSCVSPSTDECSIQELBQKLLEREGA			
	LKQLRSPFEKLASSLAYEERPRCRDELEGPEKPGNKLKQASQKSORAQOVLHLO			
	VLQLOQKRLQESLEMLKEODLLETKLRSLEGLETSPFPALEETQWEVCQKSGSIS			
	LKQOLKESQTEWNAKASILGLKQDKTRGKLEGLRTOLEGALRTKGLELRVC			
	ENELQRKNEAELLREKVNLLSQELQELRAQALARDMGPPTEPDPVALQRELERLV			
	WEEKSEKVTQYQKLOQSYVMTQRNRLKALQQLARGDSAGEPLEVDELEGADIFYE			
	DIATPEI"			
BASE COUNT	381 a	541 c	563 g	237 t
ORIGIN				
Query Match	82.9%	Score 286;	DB 9;	Length 1722;
Best Local Similarity	100.0%;	Pred. No. 96-63;		
Matches 286; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
Qy	1	TCGCAGTACAAGCTGCGCAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG		60
Db	61	TCGCAGTACAAGCTGCGCAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG		120
Qy	61	CTGCTGAGTTGGCTTCTCCACGAGACTCCGGTCAAGCAAGTCCAGTCCAAAATGGGC		120
Db	121	CTGCTGAGTTGGCTTCTCCAGAGACTCCGGTCAAGCAAGTCCAGTCCAAAATGGGC		180

QY 121 AAGAGCGAAGACTTCTTCTACATCAAGGTCCAGCAGAAAGCCGGGCTCCCATCACCCA 180
 |||||
 Db 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCCAGCAGAAAGCCGGGCTCCCATCACCCA 240
 |||||
 QY 181 GATTACACGGCACTGTCACAGCGGGGATTTAGGGGGCCAGGCTGGGTGGACTTTGACCCG 240
 |||||
 Db 241 GATTACACGGCACTGTCACAGCGGGGATTTAGGGGGCCAGGCTGGGTGGACTTTGACCCG 300
 |||||
 QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 286
 |||||
 Db 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 346
 |||||

RESULT 8
 AF123659 5492 bp mRNA linear PRI 07-APR-1999
 LOCUS Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.
 ACCESSION AF123659
 VERSION AF123659.1 GI:4572475
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5492)

AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
 Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
 TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
 and its expression is altered in multiple human tumors
 Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

JOURNAL 99199287
 MEDLINE 10097140
 PUBLISHED 2 (bases 1 to 5492)
 REFERENCE Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
 Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
 AUTHORS Direct Submission

TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
 Institute, 2338 10th street, Philadelphia, PA 19107, USA
 JOURNAL Location/Qualifiers

FEATURES
 source
 1. 5492
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8p22"
 1. 5492
 /genes="FEZ1"
 112. 1902
 /genes="FEZ1"
 /codon_start=1
 /product="FEZ1"
 /protein_id="AAD23840.1"
 /db_xref="GI:4572476"
 /translation="MGVSLSLISGHFSFKHCRASGYKLRKSHLKKLNRYSDGLLRF
 GFSGDGHGKSSKMGKSDFFYIKVQKARSHHPDYATLSSGLGGAGVDFPST
 PKLPFNSQLEKSGKAVRTAPKPLPRSGAILHSPESASHQHPAPDKPEQ
 ELKPGICSGALSDGNSNSLPTHTSSVYOLDPLVTPVGTSGEGSAHNITQIV
 LQDSNMSKALSFSGCKLGHNSKADKGCVRSPSTDECSQLEQLKLERGA
 LQKLRSEFEKLASSLAYERPRRCRDELPKPKGNLKOASQKORAOQVHLHQ
 VLQOQKRLQKLESLMKEDQLLEKLSYERKTSFPALEETQWECQKSGEIS
 LKQKQKESQTEYNKASEILGLKQKIDTRKLEGLRLETDLEALRTKLELEVC
 ENLQKKEAEALLRKNVLLLEQLQELRAQALARDMGPTTFPEVPALQRELERLR
 AELREERQGDMSGFOHERLVWKEKEKVIQYQKLOQSVYVQNRQLEKALQ
 LARGDSAGEPLEVDLEGADIPEDIAETI"
 BASE COUNT 1137 a 1704 c 1565 g 1086 t

Query Match 82.9%; Score 286; DB 9; Length 5492;
 Best Local Similarity 100.0%; Pred. No. 8,9e-63;
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAGTACAGGTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGGG 60
 |||||
 Db 172 TCGCAGTACAGGTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGGG 231
 |||||

QY 61 CTGCTGAGGTTTGGCTTCTCCAGACTCCGCTCAGCGCAAGTCCAGCTCCAAAATGGGC 120
 |||||
 Db 232 CTGCTGAGGTTTGGCTTCTCCAGACTCCGCTCAGCGCAAGTCCAGCTCCAAAATGGGC 291
 |||||
 QY 121 AAGCGCAAGACTTCTTCTACATCAAGGTCCAGCAGAAAGCCGGGCTCCCATCACCCA 180
 |||||
 Db 292 AAGCGCAAGACTTCTTCTACATCAAGGTCCAGCAGAAAGCCGGGCTCCCATCACCCA 351
 |||||
 QY 181 GATTACACGGCACTGTCACAGCGGGGATTTAGGGGGCCAGGCTGGGTGGACTTTGACCCG 240
 |||||
 Db 352 GATTACACGGCACTGTCACAGCGGGGATTTAGGGGGCCAGGCTGGGTGGACTTTGACCCG 411
 |||||

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 286
 |||||

Db 412 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 457
 |||||

RESULT 9
 AC114995 227884 bp DNA linear HTG 26-JUN-2002
 LOCUS Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered
 DEFINITION pieces.
 ACCESSION AC114995
 VERSION AC114995.3 GI:21592059
 HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 227884)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Mus musculus, clone RP23-171P5

Unpublished

2 (bases 1 to 227884)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Meneus,L.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,S., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 227884)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fato,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
 Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,

Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE

Direct Submission
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jun 26, 2002 this sequence version replaced gi:21535965.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L23473

Center clone name: 171_P_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 225162 bases at least Q40

Consensus quality: 226535 bases at least Q30

Consensus quality: 227117 bases at least Q20

Insert size: 225000; agarose-fp

Insert size: 227484; sum-of-coverage

Quality coverage: 8.6 in Q20 bases; agarose-fp

Quality coverage: 8.6 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 301: contig of 301 bp in length

* 302 401: gap of 100 bp

* 402 5831: contig of 5430 bp in length

* 5832 5931: gap of 100 bp

* 5932 173029: contig of 167098 bp in length

* 173030 173129: gap of 100 bp

* 173130 210061: contig of 36932 bp in length

* 210062 210161: gap of 100 bp

* 210162 227884: contig of 17723 bp in length.

Location/Qualifiers

1. .227884

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP23-171P5"

/clone_lib="RP23-171P5"

/clone_lib="RP23-171P5"

1. .301

/notes="assembly_fragment"

clone end:SP6

vector side:left"

402. .5831

/notes="assembly_fragment"

5932. .173029

/notes="assembly_fragment"

173130. .210061

/notes="assembly_fragment"

210162. .227884

/notes="assembly_fragment"

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

clone end:T7
vector side:right"

BASE COUNT 65784 a 48929 c 49902 g 62967 t 402 others

ORIGIN

Query Match 67.5%; Score 232.8; DB 2; Length 227884;
Best Local Similarity 84.7%; Pred. No. 3.5e-49;
Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60
|||
Db 82628 TCACAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAATCGATCTCAGATGG 82687

QY 61 CTGCTGAGGTTTGGGTTCTCCAGGACTCCGCTCAGCGCAAGTCCAGCTCCAAATGGGC 120
|||
Db 82688 CTGCTGAGGTTTGGGTTCTCCAGGACTCCGCTCAGCGCGTGGAAAGTCAAGTTCCAAATGGGA 82747

QY 121 AAGCGGAAGACTTCTTCTACATCAAGGTGAGCGCAAGAGCCGGGGTCCCATCACCCA 180
|||
Db 82748 AAAAGCGAAGACTTCTTCTACATCAAGGTGAGCGCAAGAGCCGGGGTCCCATCACCCA 82807

QY 181 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240
|||
Db 82808 GACTACACAGCCCTGTCCAGTGGGACATAGGGGGTCAAGCGGAGTAGATTTGATCCA 82867

QY 241 TCCACACCCCAAGCTCATGCTTCTCCAAATCAGTACAGTAAAGTAAAGCGGGGTCCG 300
|||
Db 82868 GCCACCCCAAGAGTCTATGCTTCTCCAAATCAGTAAAGTAAAGCGGGGTCCG 82927

QY 301 TGGCAAGG 308
|||
Db 82928 TGACAGGG 82935

RESULT 10

AC099416/c

LOCUS

DEFINITION

AC099416

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC099416

Mus musculus

SEQUENCE, 8 unordered pieces.

AC099416

AC099416.2

HTG; HTGS_PHASE1; HTGS_DRAFT.

Mus musculus

Mus musculus

Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

McPherson, J.D. and Waterston, R.H.

Unpublished

2 (bases 1 to 263546)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 263546)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

On May 23, 2002 this sequence version replaced gi:16924178.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu

----- Project Information -----

Center project name: M.BA0122M11

----- Summary Statistics -----

Sequencing vector: M13; 32%

Sequencing vector: plasmid; 68%

Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 278310 bases at least Q40
 Consensus quality: 282938 bases at least Q30
 Consensus quality: 286149 bases at least Q20
 Insert size: 250000; agarose-fp
 Insert size: 387570; sum-of-contigs
 Quality coverage: 25.74 in Q20 bases; agarose-fp
 Quality coverage: 16.86 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1204: contig of 1204 bp in length
 1205 1304: gap of unknown length
 1305 2524: contig of 1220 bp in length
 2525 2624: gap of unknown length
 2625 4136: contig of 1512 bp in length
 4137 4236: gap of unknown length
 4237 16434: contig of 12198 bp in length
 16435 16534: gap of unknown length
 16535 32106: contig of 15572 bp in length
 32107 32206: gap of unknown length
 32207 140224: contig of 108018 bp in length
 140225 140324: gap of unknown length
 140325 262693: contig of 122369 bp in length
 262694 262793: gap of unknown length
 262794 263546: contig of 753 bp in length.

FEATURES

source

1. .263546
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP23-122M11"

misc_feature 1. .1204
 /note="assembly_name:Contig127"
 misc_feature 1305. .2524
 /note="assembly_name:Contig157"
 misc_feature 2625. .4136
 /note="assembly_name:Contig360"
 misc_feature 4237. .16434
 /note="assembly_name:Contig396"
 misc_feature 16535. .32106
 /note="assembly_name:Contig397"
 misc_feature 32207. .140224
 /note="assembly_name:Contig398"
 misc_feature 140325. .262693
 /note="assembly_name:Contig399"
 misc_feature 262794. .263546
 /note="assembly_name:Contig154"
 BASE COUNT 76628 a 56149 c 55623 g 74436 t 710 others

ORIGIN

Query Match 67.5%; Score 232.8; DB 2; Length 263546;
 Best Local Similarity 84.7%; Pred. No. 3.5e-49;
 Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 1 TCAGGTACAGTGGCGAAGTCTCCACCTCAAGAGCTCAACGGTATTCGACGGG 60
 Db 252258 TCACAGTACAGTGGCGAAGTCTCCACCTCAAGAGCTCAACGGTATTCGACGGG 60
 QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCAGGCAAGTCCAGCTCCAAATGGGC 120
 Db 252198 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCAGGCAAGTCCAAATGGGA 252139
 QY 121 AAGAGCGAAGACTTCTTCTACATCAAGGTGAGCAGAAAGCCCGGGGCTCCCATCACCCA 180

Db 252138 AAAAGCGAAGACTTCTTCTACATCAAGTCAAGAGCGCGAGGCTCCCATCGCCA 252079
 QY 181 GATTACAGGCACCTCTCCAGCGGGGATTAGGGGCCAGGCTGGGGTGGACTTTACCCG 240
 Db 252078 GACTACACAGCCCTGTCCAGTGGGACATAGGGGTGACAGCGGAGTAGATTTGATCCA 252019
 QY 241 TCCACACCCCAAGCTCATGCCCTTCTCCAACTAGTCTAGTAATGCTAAGCGGGGTCCG 300
 Db 252018 GCCACCCCAAGACTCATGCCCTTCTCCAACTAGTCTAGTAATGCTAAGCGGGGTCT 251959
 QY 301 TGGCAAGG 308
 Db 251958 TGACAGGG 251951

RESULT 11

AC108987

LOCUS

AC108987 191210 bp DNA linear HTG 13-JUL-2002
 Rattus norvegicus clone CH230-115K1, *** SEQUENCING IN PROGRESS

DEFINITION *** 78 unordered pieces.

ACCESSION AC108987.3 GI:21737647

VERSION AC108987.3 HTG; HTGS PHASE1.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 191210)

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Mashey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Roife, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 191210)

AUTHORS Worley, K.C.

TITLE
JOURNAL

Direct Submission
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 191210)
Worley, K.C.

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18846600.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPMW
Center clone name: CH230-115K1
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115719 bases at least Q40
Consensus quality: 122885 bases at least Q30
Consensus quality: 129025 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 78 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1041: contig of 1041 bp in length
* 1042
* 1141: gap of unknown length
* 1142
* 2262: contig of 1121 bp in length
* 2263
* 2382: gap of unknown length
* 2383
* 3538: contig of 1176 bp in length
* 3539
* 3638: gap of unknown length
* 3639
* 4775: contig of 1137 bp in length
* 4776
* 4875: gap of unknown length
* 4876
* 6050: contig of 1175 bp in length
* 6051
* 6150: gap of unknown length
* 6151
* 7660: contig of 1510 bp in length
* 7661
* 9331: contig of 1571 bp in length
* 9332
* 9432
* 11027: contig of 1596 bp in length
* 11028
* 11279: gap of unknown length
* 11128
* 12769: contig of 1642 bp in length
* 12770
* 12869: gap of unknown length
* 14092: contig of 1223 bp in length
* 14093
* 14192: gap of unknown length
* 14193
* 15228: contig of 1036 bp in length
* 15229
* 15328: gap of unknown length
* 15329
* 16709: contig of 1381 bp in length
* 16710
* 16809: gap of unknown length
* 16810
* 18195: contig of 1386 bp in length
* 18196
* 18295: gap of unknown length
* 18296
* 19412: contig of 1117 bp in length
* 19413
* 19512: gap of unknown length
* 20959: contig of 1447 bp in length
* 20960
* 21059: gap of unknown length
* 21060
* 21161: contig of 1102 bp in length
* 21162
* 22261: gap of unknown length
* 22262
* 23587: contig of 1326 bp in length
* 23588
* 23687: gap of unknown length
* 25231: contig of 1544 bp in length
* 25232
* 25331: gap of unknown length

25332
* 27136
* 27136
* 27175: contig of 1804 bp in length
* 27176
* 29275: gap of unknown length
* 29276
* 30846: contig of 1571 bp in length
* 30847
* 30946: gap of unknown length
* 30947
* 32097: contig of 1144 bp in length
* 32098
* 32190: gap of unknown length
* 32191
* 34039: contig of 1849 bp in length
* 34040
* 34139: gap of unknown length
* 34140
* 35408: contig of 1269 bp in length
* 35409
* 35508: gap of unknown length
* 35509
* 37018: contig of 1510 bp in length
* 37019
* 37118: gap of unknown length
* 37119
* 38260: contig of 1142 bp in length
* 38261
* 38360: gap of unknown length
* 38361
* 40083: contig of 1723 bp in length
* 40084
* 40183: gap of unknown length
* 40184
* 41200: contig of 1017 bp in length
* 41201
* 41300: gap of unknown length
* 41301
* 43365: contig of 2065 bp in length
* 43366
* 43465: gap of unknown length
* 43466
* 45443: contig of 1878 bp in length
* 45444
* 46987: gap of unknown length
* 46988
* 47087: contig of 1544 bp in length
* 47088
* 49404: contig of 2317 bp in length
* 49405
* 49504: gap of unknown length
* 49505
* 50823: contig of 1319 bp in length
* 50824
* 50923: gap of unknown length
* 52451: contig of 1528 bp in length
* 52452
* 52551: gap of unknown length
* 52552
* 54411: contig of 1890 bp in length
* 54412
* 54541: gap of unknown length
* 54542
* 56442: contig of 1901 bp in length
* 56443
* 56542: gap of unknown length
* 56543
* 57499: contig of 1207 bp in length
* 57500
* 57849: gap of unknown length
* 57850
* 60138: contig of 2289 bp in length
* 60139
* 60238: gap of unknown length
* 60239
* 62467: contig of 2229 bp in length
* 62468
* 62567: gap of unknown length
* 64364: contig of 1797 bp in length
* 64365
* 64464: gap of unknown length
* 64465
* 66889: contig of 2425 bp in length
* 66890
* 66989: gap of unknown length
* 68300: contig of 1841 bp in length
* 68301
* 68930: gap of unknown length
* 68931
* 71227: contig of 2296 bp in length
* 71228
* 71326: gap of unknown length
* 71327
* 73127: contig of 2386 bp in length
* 73128
* 73812: gap of unknown length
* 73813
* 76871: contig of 3059 bp in length
* 76872
* 76971: gap of unknown length
* 76972
* 79844: contig of 2873 bp in length
* 79845
* 79944: gap of unknown length
* 79945
* 81259: contig of 1315 bp in length
* 81260
* 81359: gap of unknown length
* 81360
* 83690: contig of 2331 bp in length
* 83691
* 83790: gap of unknown length
* 83791
* 86195: contig of 2405 bp in length
* 86196
* 86295: gap of unknown length
* 86296
* 88065: contig of 1770 bp in length
* 88066
* 88166
* 90384: contig of 2219 bp in length
* 90385
* 90484: gap of unknown length
* 90485
* 91999: contig of 1515 bp in length
* 92000
* 92099: gap of unknown length
* 92100
* 94583: contig of 2484 bp in length
* 94584
* 94683: gap of unknown length
* 94684

Query Match

66.9%; Score 230.8; DB 2; Length 191210;

Db 46437 GCCCGCCTGACCCCAAGTCAAGCGCTCGGCCAGGACTCGTTCAAGGGCATGGCGT 46378
QY 72 TGGCTTCTCCAGGACTCCGGTCAAGGAGTTCAGTCCAAAATGGGCAAGAGCAAGA 131
Db 46377 GGACTTCGGCGACATCGAGCGGACGGCATGTTTCGACCTGTACGTGCGCAACATCACCAC 46318
QY 132 CTTCTTCTACATCAAGTTCAGCAGAAAGCCCGGGCTCCCATCACCCAGATTACACGCG 191
Db 46317 CTCCTTCGGCATCCAGGAGAGCAACTTCGCCCTTCGTCAACACCGCGCGCACCGCGCG 46258
QY 192 ACTGTCACGCGGGGATTTAGGGGGCCAGGCTGGGTGGAC 231
Db 46257 GCTGCGCGCGGCTCTGGCGCGGAGGCGCGCTGGCAC 46218

RESULT 14
AR103043
LOCUS AR103043 1981 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 3 from patent US 6087147.
ACCESSION AR103043
VERSION AR103043.1 GI:12814631
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Ito,Y.
TITLE .alpha.-amylase gene having ability for highly producing
maltopentaose, vector containing said gene and transformant
JOURNAL Patent: US 6087147-A 3 11-JUL-2000;
FEATURES Location/Qualifiers
source 1..1981
BASE COUNT 342 a 678 c 678 g 283 t
ORIGIN

Query Match 11.5%; Score 39.6; DB 6; Length 1981;
Best Local Similarity 52.4%; Pred. No. 9.8; Mismatches 79; Indels 0; Gaps 0;
Matches 87; Conservative 0;
QY 3 GCAGTACAAGCTGCGCAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGGCT 62
Db 792 GCAGGTCACGGTGACCGAGTTCGCTACGCGAAGGAGCTCTACGGCAAGTTCGCGCGGG 851
QY 63 GCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGAGTTCAGCTCCAAAATGGGCAA 122
Db 852 CGCAAGCTGCGCGACCTGCAGACCTTCGCGCCCGCAGCTGGAACCTGATGCCCGCAGCAA 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGTTCAGCGCAAGGAGCGCGGGGC 168
Db 912 GGCCATCGCTTTCGTGCAACCAACGACGAGCGCGCGCCACGGC 957

RESULT 15
AR103044
LOCUS AR103044 1981 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 4 from patent US 6087147.
ACCESSION AR103044
VERSION AR103044.1 GI:12814632
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Ito,Y.
TITLE .alpha.-amylase gene having ability for highly producing
maltopentaose, vector containing said gene and transformant
JOURNAL Patent: US 6087147-A 4 11-JUL-2000;
FEATURES Location/Qualifiers
source 1..1981
BASE COUNT 343 a 679 c 678 g 281 t

ORIGIN

Query Match 11.5%; Score 39.6; DB 6; Length 1981;
Best Local Similarity 52.4%; Pred. No. 9.8; Mismatches 79; Indels 0; Gaps 0;
Matches 87; Conservative 0;
QY 3 GCAGTACAAGCTGCGCAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGGCT 62
Db 792 GCAGGTCACGGTGACCGAGTTCGCTACGCGAAGGAGCTCTACGGCAAGTTCGCGCGGG 851
QY 63 GCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGAGTTCAGCTCCAAAATGGGCAA 122
Db 852 CGCAAGCTGCGCGACCTGCAGACCTTCGCGCCCGCAGCTGGAACCTGATGCCCGCAGCAA 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGTTCAGCGCAAGGAGCGCGGGGC 168
Db 912 GGCCATCGCTTTCGTGCAACCAACGACGAGCGCGCGCCACGGC 957

Search completed: June 15, 2003, 02:18:31
Job time : 901.829 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:38:41 ; Search time 79.43 Seconds
(without alignments)

9781.434 Million cell updates/sec

Title: US-09-513-888C-1_COPY_112_456

Perfect score: 345

Sequence: 1 tcgcagtcacagtcgcgcaaa.....gcaggagaaagcaaatggg 345

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 2185239 seqs, 1125995159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N_Geneseq_101002.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 25: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	100.0	9048	21	AAA64507
2	286	82.9	633	21	AAA64511
3	286	82.9	1512	21	AAA64513
4	286	82.9	1614	21	AAA64512
5	286	82.9	1722	21	AAA64515
6	286	82.9	1791	21	AAA64509
7	286	82.9	5432	21	AAA64508
8	284.4	82.4	1692	21	AAA64514
9	115	33.3	2411	22	AA530637

10	115	33.3	2411	22	AA528699	Genomic sequence #
11	106	30.7	404	21	AAA64510	Nucleotide sequenc
12	39.6	11.5	1981	20	AA537293	Pseudomonas alpha-
13	39.6	11.5	1981	20	AA537294	Pseudomonas alpha-
14	39.6	11.5	1981	20	AA537295	Pseudomonas alpha-
15	39.6	11.5	1981	20	AA537296	Pseudomonas alpha-
16	39.6	11.5	1981	20	AA537297	Pseudomonas alpha-
17	39.6	11.5	1981	20	AA537298	Pseudomonas alpha-
18	39.6	11.5	3496	12	AAQ14465	Desulfotomacrose synt
19	38.6	11.2	324	19	AAV28847	Maltotetraose synth
20	38.6	11.2	324	21	AAZ40359	Wild type cytochrome
21	38.6	11.2	873	21	AA98361	D. vulgaris cytochro
22	38.6	11.2	873	21	AA98362	D. vulgaris cytochro
23	38.6	11.2	873	21	AA98363	Cytochrome C3 gene
24	38.6	11.2	873	21	AA98364	Desulfotomacrose synt
25	37	10.7	873	22	AA164981	Desulfotomacrose synt
26	36.4	10.6	1915	19	AA60319	DNA sequence of th
27	36.4	10.6	2267	18	AA63355	Granule bound star
28	36.4	10.6	4800	19	AAV29752	Zea mays waxy gene
29	36	10.4	560	22	AA530595	DNA encoding novel
30	36	10.4	560	22	AA528127	Novel cDNA encodin
31	35.2	10.2	362	21	AA531932	Plant microsatelli
32	35	10.1	11115	23	ABL50562	Micromonospora car
33	34.6	10.0	277	24	ABN23622	Human ORFX polynuc
34	34.6	10.0	5267	19	AAV32449	Human receptor tyr
35	34.4	10.0	413	22	AA192339	Human polynucleoti
36	34.4	10.0	2681	17	AA130870	Engineered 95 kD p
37	34	9.9	1665	23	AA565391	DNA encoding novel
38	34	9.9	13842	21	AAZ87297	S. venezuelae macr
39	34	9.9	36778	21	AAZ87318	S. venezuelae pik
40	34	9.9	37948	21	AAZ87285	S. venezuelae pik
41	34	9.9	38506	21	AAZ75633	Nucleotide sequenc
42	34	9.9	38506	21	AAZ56001	Recombinant cosmid
43	34	9.9	43280	18	AAZ80413	Tylosin synthase
44	33.8	9.8	4114	23	ABL05659	Drosophila melanog
45	33.8	9.8	9244	23	ABL05658	Drosophila melanog

ALIGNMENTS

RESULT 1
AAA64507
ID AAA64507 standard; DNA; 9048 BP.

AC AAA64507;

DT 02-JAN-2001 (first entry)

DE Nucleotide sequence comprising the human FEZ1 gene.

Human: FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
tumour proliferation; tubulin; microtubule; protein E1-gamma;
tubulin polymerisation disorder; mitosis initiation; cell proliferation;
cell growth; cell shape; cell rigidity; cell motility; DNA replication;
tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.

XX WO2000050565-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Ishii H;

XX WPI; 2000-558396/51.

XX


```

RESULT 3
AAA64513
ID AAA64513 standard; cDNA; 1512 BP.
XX
AC AAA64513;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript D14.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 1..1512
FT FT /*tag= a
FT FT /product= "truncated FEZ1"
XX
PN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Ishii H;
XX
DR WPI; 2000-558396/51.
DR P-PSDB; AAB08720.
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
PS Disclosure; Fig 5F; 255pp; English.
XX
CC The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC Efi-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;
Query Match 82.9%; Score 286; DB 21; Length 1512;
Best Local Similarity. 100.0%; Pred. No. 1.4e-71;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TCGCAGTACAAGCTGGCGAAGTCTCCCACTCAAGAGCTCAACCGGTATTCCGACGGG 60
|||||

```

QY

CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of the
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX

XX Sequence 1614 BP; 367 A; 500 C; 522 G; 225 T; 0 other;

Query Match 82.9%; Score 286; DB 21; Length 1614;

Best Local Similarity 100.0%; Pred. No. 1.4e-71;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 60

|||||

Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120

|||||

QY 61 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120

|||||

Db 121 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 180

|||||

QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 180

|||||

Db 181 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 240

|||||

QY 181 GATTACACGGCACTGCCAGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 240

|||||

Db 241 GATTACACGGCACTGCCAGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 300

|||||

QY 241 TCCACACCCCGCAAGCTCATGTCCTTCTCCATCAGCTAGAAATGG 286

|||||

Db 301 TCCACACCCCGCAAGCTCATGTCCTTCTCCATCAGCTAGAAATGG 346

RESULT 5

AAA64515
ID AAA64515 standard; cDNA; 1722 BP.

XX CDS

AC AAA64515;

XX 02-JAN-2001 (first entry)

XX Nucleotide sequence of truncated FEZ1 transcript G3612.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
XX tumour proliferation; tubulin; microtubule; protein Efi-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1722

XX FT /*tag= a

XX FT /product= "truncated FEZ1"

XX PN WQ20000565-A2.

XX PD 31-AUG-2000.

XX PF 25-FEB-2000; 2000MO-US04950.

XX PR 25-FEB-1999; 99US-0121537.

XX PA (UJJE-) UNIV JEFFERSON THOMAS.

XX PI Croce CM, Ishii H;

XX DR WPI; 2000-558396/51.

DR P-PSDB; AAB08722.

XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
XX cancer -

PS Disclosure; Fig 5H; 255pp; English.

XX The present sequence encodes a truncated human FEZ1 polypeptide. The

CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour

CC suppressor gene, located at chromosome location 8p22. Decreased

CC or no expression of FEZ1 is detected in a variety of cancer cells.

CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1

CC also interacts with tubulin, with microtubules, and with protein

CC Efi-gamma. Post-translational phosphorylation and dephosphorylation

CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene

CC expression are useful for inducing cells to proliferate. Compounds

CC which modulate FEZ1 association with tubulin are useful for alleviating

CC tubulin hyper- or hypo- polymerisation disorders, such as those

CC associated with aberrant initiation of mitosis, modulation of the

CC initiation and rate of cell proliferation and cell growth, modulation of

CC cell shape, cell rigidity, cell motility, rate and stage of cellular

CC DNA replication, intracellular distribution of organelles, metastatic

CC potential of cell and cellular transformation from a non-cancerous to

CC cancerous phenotype. Compounds which modulate FEZ1 binding and

CC phosphorylation are also useful for alleviating a disorder, such as

CC tumorigenesis, tumour survival, growth and metastasis.

XX

SQ Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;

Query Match 82.9%; Score 286; DB 21; Length 1722;

Best Local Similarity 100.0%; Pred. No. 1.4e-71;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 60

|||||

Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120

|||||

QY 61 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120

|||||

Db 121 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 180

|||||

QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 180

|||||

Db 181 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 240

|||||

QY 181 GATTACACGGCACTGCCAGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 240

|||||

Db 241 GATTACACGGCACTGCCAGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 300

|||||

QY 241 TCCACACCCCGCAAGCTCATGTCCTTCTCCATCAGCTAGAAATGG 286

|||||

Db 301 TCCACACCCCGCAAGCTCATGTCCTTCTCCATCAGCTAGAAATGG 346

RESULT 6

AAA64509

ID AAA64509 standard; cDNA; 1791 BP.

XX AC AAA64509;

XX 02-JAN-2001 (first entry)

XX cDNA sequence encoding a human FEZ1 polypeptide.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.

XX Homo sapiens.

```

FH Key Location/Qualifiers
FT CDS 1..1791
FT FT /*tag= a
FT FT /product= "FEZ1"
XX
XX
XX WO2000050565-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US04950.
XX
XX 25-FEB-1999; 99US-0121537.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Croce CM, Ishii H;
XX
XX WPI; 2000-558396/51.
XX
XX P-PSDB; AAB08715.
XX
XX New polynucleotide homologous with a portion of one strand of the human
XX FEZ1 gene, useful for alleviating abnormal cell proliferation such as
XX cancer -
XX
XX Claim 7; Fig 5I; 255pp; English.
XX
XX The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a
XX tumour suppressor gene, located at chromosome location 8p22. Decreased
XX or no expression of FEZ1 is detected in a variety of cancer cells.
XX Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
XX also interacts with tubulin, with microtubules, and with protein
XX BFI-gamma. Post-translational phosphorylation and dephosphorylation
XX modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
XX expression are useful for inducing cells to proliferate. Compounds
XX which modulate FEZ1 association with tubulin are useful for alleviating
XX tubulin hyper- or hypo- polymerisation disorders, such as those
XX associated with aberrant initiation of mitosis, modulation of the
XX initiation and rate of cell proliferation and cell growth, modulation of
XX cell shape, cell rigidity, cell motility, rate and stage of cellular
XX DNA replication, intracellular distribution of organelles, metastatic
XX potential of cell and cellular transformation from a non-cancerous to
XX cancerous phenotype. Compounds which modulate FEZ1 binding and
XX phosphorylation are also useful for alleviating a disorder, such as
XX tumorigenesis, tumour survival, growth and metastasis.
XX
XX Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;
XX
XX
XX Query Match 82.9%; Score 286; DB 21; Length 1791;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-71;
XX Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
DB 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120
QY 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAAATGGGC 120
DB 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAAATGGGC 180
QY 121 AGAGCGGAAGACTTCTTACATCAAGGTCAGCCAGAAAGCCGGGGTCCCATCACCCA 180
DB 181 AGAGCGGAAGACTTCTTACATCAAGGTCAGCCAGAAAGCCGGGGTCCCATCACCCA 240
QY 181 GATTACACGCACTGTCCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240
DB 241 GATTACACGCACTGTCCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300
QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286
DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346
XX
XX RESULT 7

```

AAA64508

ID AAA64508 standard; cDNA; 5492 BP.

XX

AC AAA64508;

XX

DT 02-JAN-2001 (first entry)

XX

DE cDNA sequence of the wild type human FEZ1 gene.

XX

Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth; tumour proliferation; tubulin; microtubule; protein BFI-gamma; tubulin polymerisation disorder; mitosis initiation; cell proliferation; cell growth; cell shape; cell rigidity; cell motility; DNA replication; tumorigenesis; tumour survival; metastasis; ss.

XX

OS Homo sapiens.

XX

XX WO2000050565-A2.

XX

XX 31-AUG-2000.

XX

XX 25-FEB-2000; 2000WO-US04950.

XX

XX 25-FEB-1999; 99US-0121537.

XX

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX

XX Croce CM, Ishii H;

XX

XX WPI; 2000-558396/51.

XX

New polynucleotide homologous with a portion of one strand of the human FEZ1 gene, useful for alleviating abnormal cell proliferation such as cancer -

XX

XX Example 2; Fig 5B; 255pp; English.

XX

The present sequence represents the cDNA sequence of the human FEZ1 gene. FEZ1 is a tumour suppressor gene, located at chromosome location 8p22. Decreased or no expression of FEZ1 is detected in a variety of cancer cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts with tubulin, with microtubules, and with protein BFI-gamma. Post-translational phosphorylation and dephosphorylation modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are useful for inducing cells to proliferate. Compounds which modulate FEZ1 association with tubulin are useful for alleviating tubulin hyper- or hypo- polymerisation disorders, such as those associated with aberrant initiation of mitosis, modulation of the initiation and rate of cell proliferation and cell growth, modulation of cell shape, cell rigidity, cell motility, rate and stage of cellular DNA replication, intracellular distribution of organelles, metastatic potential of cell and cellular transformation from a non-cancerous to cancerous phenotype. Compounds which modulate FEZ1 binding and phosphorylation are also useful for alleviating a disorder, such as tumorigenesis, tumour survival, growth and metastasis.

Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;

Query Match

Best Local Similarity 82.9%; Score 286; DB 21; Length 5492;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60

DB

172 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 231

QY

61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAAATGGGC 120

DB

232 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAAATGGGC 291

QY

121 AGAGCGGAAGACTTCTTACATCAAGGTCAGCCAGAAAGCCGGGGTCCCATCACCCA 180

DB

292 AAGAGCGAAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCGGGGTCCCATCACCCA 351

QY 181 GATTACACGGCACTGTCCAGCGGGGATTAGGGGGCCAGGCTGGGGTGAGCTTTGACCCG 240
DB 352 GATTACACGGCACTGTCCAGCGGGGATTAGGGGGCCAGGCTGGGGTGAGCTTTGACCCG 411
QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCATCAGCTAGAAATGG 286
DB 412 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCATCAGCTAGAAATGG 457

RESULT 8

AA64514

ID AAA64514 standard; cDNA; 1692 BP.

AC

AAA64514;

XX 02-JAN-2001 (first entry)

DE Nucleotide sequence of truncated FEZ1 transcript G3611.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EFL-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 1..1692

FT CDS

FT /*tag= a

FT /*product= "truncated FEZ1"

XX WO2000050565-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Ishii H;

PI WPI; 2000-558396/51.

DR P-PSDB; AAB08721.

XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -

XX Disclosure; Fig 5G; 255pp; English.

XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EFL-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.

XX SQ Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;

Query Match 82.4%; Score 284.4; DB 21; Length 1692;

Best Local Similarity 99.7%; Pred. No. 4e-71;

Matches 285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTTCCACGGG 60
DB 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTTCCACGGG 120

QY 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCCGTACGGCAAGTCCAGCTCCAAAATGGGC 120
DB 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCCGTACGGCAAGTCCAGCTCCAAAATGGGC 180

QY 121 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGCCAGAAAGCCGGGCTCCCATCACCCA 180
DB 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGCCAGAAAGCCGGGCTCCCATCACCCA 240

QY 181 GATTACACGGCACTGTCCAGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 240
DB 241 GATTACACGGCACTGTCCAGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCATCAGCTAGAAATGG 286
DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCATCAGCTAGAAATGG 346

RESULT 9

AAS30637

ID AAS30637 standard; DNA; 2411 BP.

XX AAS30637;

XX 21-NOV-2001 (first entry)

XX DNA encoding novel lung cancer antigen, Seq ID No 89.

XX Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasia;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; skin aging;
KW ocular disorder; wound healing; organ transplantation; ds.

XX Homo sapiens.

XX WO200155300-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01238.

XX 31-JAN-2000; 2000US-0179065.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465565/50.

XX Isolated nucleic acid molecule encoding a lung cancer antigen is used
PT in preventing, treating or ameliorating a medical condition -

XX Disclosure; SEQ ID No 89; 475pp; English.

XX The invention relates to novel isolated lung cancer antigen

CC polynucleotides (I) and polypeptides (II). (I) and (II) are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are
CC also used in diagnosing a pathological condition or susceptibility to a
CC pathological condition, in particular, lung cancer. The antibodies to

CC (II) can also be used in alleviating symptoms associated with the
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC treated include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. The polypeptides can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. AAS30580-AAS30685 represent novel human lung cancer antigen
CC coding sequences, PCR primers and related sequences of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2411 BP; 558 A; 681 C; 733 G; 439 T; 0 other;

Query Match 33.3%; Score 115; DB 22; Length 2411;
Best Local Similarity 100.0%; Pred. No. 7.7e-23;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGGTAAG 290
Db 1 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGGTAAG 60
QY 291 CGGGGTCGCTGCGCAAGGTAAGTGGTTGGAACGACGAGAAAGCAAAATGGG 345
Db 61 CGGGGTCGCTGCGCAAGGTAAGTGGTTGGAACGACGAGAAAGCAAAATGGG 115

RESULT 10

AAS28699

ID AAS28699 standard; DNA; 2411 BP.

AC AAS28699;

DT 07-NOV-2001 (first entry)

DE Genomic sequence #539 encoding for novel human respiratory antigen.

KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active; ds.

XX Homo sapiens.

XX WO200155448-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01333.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-476224/51.
 Isolated polypeptide for treating, preventing and/or diagnosing
 disorders related to the respiratory system including respiratory
 cancers and also for testing and detection e.g. diagnosis -
 Disclosure; SED ID No 1133; 546pp; English.
 The present invention relates to the isolation of novel human
 respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
 sequences encoding for these polypeptides. The sequences of the
 invention are useful for preventing, treating and/or diagnosing
 disorders related to the respiratory system including throat
 disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
 lung disorders (e.g. pneumonia, allergic disorders e.g. asthma,
 pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
 the respiratory tissues e.g. lung cancer. The polynucleotide
 sequences of the invention are useful in gene therapy and antisense
 therapy.
 AAS28161-AAS28764 represent genomic sequences encoding for novel
 human respiratory antigens.
 Note: The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2411 BP; 558 A; 581 C; 733 G; 439 T; 0 other;
 SQ Query Match 33.3%; Score 115; DB 22; Length 2411;
 Best Local Similarity 100.0%; Pred. NO. 7.7e-23;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 231 CTTTGACCCGTCACACCCCAAGCTCATGCTTCTTCCAAATCAGCTAGAAATGGTAAG 290
 DB 1 CTTTGACCCGTCACACCCCAAGCTCATGCTTCTTCCAAATCAGCTAGAAATGGTAAG 60
 QY 291 CGGGGGTTCGTCGCAAGGTAAGTGGTGGTGAACGAGGAGAAAGCAAAATGGG 345
 DB 61 CGGGGGTTCGTCGCAAGGTAAGTGGTGGTGAACGAGGAGAAAGCAAAATGGG 115
 RESULT 11
 AAA64510
 ID AAA64510 standard; cDNA; 404 BP.
 XX
 AC AAA64510;
 XX
 DT 02-JAN-2001 (first entry)
 XX
 DE Nucleotide sequence of truncated FEZ1 transcript E1678.
 XX
 KW Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;
 tumour proliferation; tubulin; microtubule; protein E1-gamma;
 tubulin polymerisation disorder; mitosis initiation; cell proliferation;
 cell growth; cell shape; cell rigidity; cell motility; DNA replication;
 tumorigenesis; tumour survival; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..231
 FT /tag= a
 FT /product= "truncated FEZ1"
 FT
 XX WO2000050565-A2.
 XX
 PD 31-AUG-2000.
 XX
 PF 25-FEB-2000; 2000WO-US04950.
 XX
 PR 25-FEB-1999; 99US-0121537.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Croce CM, Ishii H;
 XX
 DR WPI; 2000-558396/51.
 DR P-PSDB; AAB08717.
 XX
 PT New polynucleotide homologous with a portion of one strand of the human
 FEZ1 gene, useful for alleviating abnormal cell proliferation such as
 cancer -
 XX
 PS Disclosure; Fig 5C; 255pp; English.
 XX
 CC The present sequence encodes a truncated human FEZ1 polypeptide. The
 encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
 suppressor gene, located at chromosome location 9p22. Decreased
 or no expression of FEZ1 is detected in a variety of cancer cells.
 CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
 also interacts with tubulin, with microtubules, and with protein
 E1-gamma. Post-translational phosphorylation and dephosphorylation
 CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
 expression are useful for inducing cells to proliferate. Compounds
 CC which modulate FEZ1 association with tubulin are useful for alleviating
 CC tubulin hyper- or hypo- polymerisation disorders, such as those
 CC associated with aberrant initiation of mitosis, modulation of the

CC initiation and rate of cell proliferation and cell growth, modulation of
 CC cell shape, cell rigidity, cell motility, rate and stage of cellular
 CC DNA replication, intracellular distribution of organelles, metastatic
 CC potential of cell and cellular transformation from a non-cancerous to
 CC cancerous phenotype. Compounds which modulate FEZ1 binding and
 CC phosphorylation are also useful for alleviating a disorder, such as
 CC tumorigenesis, tumour survival, growth and metastasis.
 XX
 SQ Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;

Query Match 30.7%; Score 106; DB 21; Length 404;
 Best Local Similarity 95.6%; Pred. No. 1.7e-20;
 Matches 109; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 60
 DB 61 TCGCAGTACAAGCTCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 120
 QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCCGGTCAAGCAAGTCCAGTCCAAA 114
 DB 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCCGGTCAAGCAAGTCCAGTCCAAA 174

RESULT 12

AAX37293
 ID AAX37293 standard; DNA; 1981 BP.

XX
 AC AAX37293;

XX
 DT 27-JUL-1999 (first entry)

XX Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6116).

XX Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.

XX Pseudomonas sp.

XX Synthetic.

XX JP11123081-A.

XX PD 11-MAY-1999.

XX PF 21-OCT-1997; 97JP-0305071.

XX PR 21-OCT-1997; 97JP-0305071.

XX PA (NORQ) NORINSUISANSHO SHOKUHIN SOGO.

XX WPI; 1999-340517/29.

XX DR P-PSDB; AAY17904.

XX PT New alpha-amylase gene - useful for production of malto-pentaose

XX Claim 4; Page 6-8; 23pp; Japanese.

XX Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase
 CC variants (AAY17904-Y17909). The variants are derived from the
 CC Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.
 CC et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants
 CC have a replacement of the 57th or the 130th amino acid residue in the
 CC sequence as compared to the wild-type. The sequences are deposited under
 CC the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The
 CC variants can be used to transform plants to produce maltopentaose. The
 CC enzymes prepared from the alpha-amylase variants produce maltopentaose
 CC in a high yield.

XX SQ Sequence 1981 BP; 342 A; 678 C; 678 G; 283 T; 0 other;

Query Match 11.5%; Score 39.6; DB 20; Length 1981;
 Best Local Similarity 52.4%; Pred. No. 0.22;
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 3 GCAGTACAAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGGCT 62

DB 792 GCAGGTCAAGTACCGGAGTTCGGCTACGGCAAGAGCTCTACGGCAAGTTCGGCGGGG 851
 QY 63 GCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGTCCAGTCCAAATGGCAA 122
 DB 852 CGGCAAGCTGGCGACCTTCGGCCCGGAGCTCTACGGCAAGTTCGGCGGGG 911
 QY 123 GAGCAAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCCGGGG 168
 DB 912 GGCCATCGCTTTCGTCGACCAACGACGAGCGGCGGCACGGC 957

RESULT 13

AAX37294

ID AAX37294 standard; DNA; 1981 BP.

XX
 AC AAX37294;

XX 27-JUL-1999 (first entry)

XX Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6119).

XX Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.

XX Pseudomonas sp.

XX Synthetic.

XX JP11123081-A.

XX PD 11-MAY-1999.

XX PF 21-OCT-1997; 97JP-0305071.

XX PR 21-OCT-1997; 97JP-0305071.

XX PA (NORQ) NORINSUISANSHO SHOKUHIN SOGO.

XX WPI; 1999-340517/29.

XX DR P-PSDB; AAY17905.

XX PT New alpha-amylase gene - useful for production of malto-pentaose

XX Claim 5; Page 9-11; 23pp; Japanese.

XX Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase
 CC variants (AAY17904-Y17909). The variants are derived from the
 CC Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.
 CC et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants
 CC have a replacement of the 57th or the 130th amino acid residue in the
 CC sequence as compared to the wild-type. The sequences are deposited under
 CC the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The
 CC variants can be used to transform plants to produce maltopentaose. The
 CC enzymes prepared from the alpha-amylase variants produce maltopentaose
 CC in a high yield.

XX SQ Sequence 1981 BP; 343 A; 679 C; 678 G; 281 T; 0 other;

Query Match 11.5%; Score 39.6; DB 20; Length 1981;
 Best Local Similarity 52.4%; Pred. No. 0.22;
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 3 GCAGTACAAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGGCT 62

DB 792 GCAGGTCAAGTACCGGAGTTCGGCTACGGCAAGAGCTCTACGGCAAGTTCGGCGGGG 851

QY 63 GCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGTCCAGTCCAAATGGCAA 122

DB 852 CGGCAAGCTGGCGACCTTCGGCCCGGAGCTCTACGGCAAGTTCGGCGGGG 911

QY 123 GAGCAGAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCCGGGG 168

DB 912 GGCCATCGCTTTCGTCGACCAACGACGAGCGGCGGCACGGC 957

RESULT 14

AA37295
ID AAX37295 standard; DNA; 1981 BP.
XX AC AAX37295;
XX DT 27-JUL-1999 (first entry)
XX DE Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6115).
XX KW Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.
XX OS Pseudomonas sp.
XX OS Synthetic.
XX PN JP11123081-A.
XX PD 11-MAY-1999.
XX PF 21-OCT-1997; 97JP-0305071.
XX PR 21-OCT-1997; 97JP-0305071.
XX PA (NORQ) NORINSUISANSHO SHOKUHN SOGO.
XX WI; 1999-340517/29.
XX DR P-PSDB; AAY17906.
XX PT New alpha-amylase gene - useful for production of malto-pentaose
XX PS Claim 6; Page 11-14; 23pp; Japanese.
XX CC Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase
variants (AAY17904-Y17909). The variants are derived from the
Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.
et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants
have a replacement of the 57th or the 130th amino acid residue in the
sequence as compared to the wild-type. The sequences are deposited under
the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The
variants can be used to transform plants to produce maltopentaose. The
enzymes prepared from the alpha-amylase variants produce maltopentaose
in a high yield.

XX SQ Sequence 1981 BP; 342 A; 679 C; 678 G; 282 T; 0 other;
Query Match 11.5%; Score 39.6; DB 20; Length 1981;
Best Local Similarity 52.4%; Pred. No. 0.22;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GCAGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
Db 792 GCAGGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851
QY 63 GCTGAGGTTGGCTTCTCCAGACTCCGCTCAGCGCAAGTCCAGTCCAAATGGCAA 122
Db 852 CGGCAAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAAGGTCAAGGTCAAGGTCAAGGTCAAGGT 168
Db 912 GGCCATCGCTTTCGTGCAACACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 957

RESULT 15

AA37296
ID AAX37296 standard; DNA; 1981 BP.
XX AC AAX37296;
XX DT 27-JUL-1999 (first entry)
XX DE Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6117).
XX OS Pseudomonas sp.
XX OS Synthetic.
XX PN JP11123081-A.
XX PD 11-MAY-1999.
XX PF 21-OCT-1997; 97JP-0305071.
XX PR 21-OCT-1997; 97JP-0305071.
XX PA (NORQ) NORINSUISANSHO SHOKUHN SOGO.
XX WI; 1999-340517/29.
XX DR P-PSDB; AAY17906.
XX PT New alpha-amylase gene - useful for production of malto-pentaose
XX PS Claim 6; Page 11-14; 23pp; Japanese.
XX CC Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase
variants (AAY17904-Y17909). The variants are derived from the
Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.
et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants
have a replacement of the 57th or the 130th amino acid residue in the
sequence as compared to the wild-type. The sequences are deposited under
the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The
variants can be used to transform plants to produce maltopentaose. The
enzymes prepared from the alpha-amylase variants produce maltopentaose
in a high yield.

XX SQ Sequence 1981 BP; 342 A; 679 C; 678 G; 282 T; 0 other;
Query Match 11.5%; Score 39.6; DB 20; Length 1981;
Best Local Similarity 52.4%; Pred. No. 0.22;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GCAGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
Db 792 GCAGGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851
QY 63 GCTGAGGTTGGCTTCTCCAGACTCCGCTCAGCGCAAGTCCAGTCCAAATGGCAA 122
Db 852 CGGCAAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAAGGTCAAGGTCAAGGTCAAGGTCAAGGT 168
Db 912 GGCCATCGCTTTCGTGCAACACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 957

Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.

Pseudomonas sp.

Synthetic.

JP11123081-A.

11-MAY-1999.

21-OCT-1997; 97JP-0305071.

21-OCT-1997; 97JP-0305071.

(NORQ) NORINSUISANSHO SHOKUHN SOGO.

WI; 1999-340517/29.

P-PSDB; AAY17907.

New alpha-amylase gene - useful for production of malto-pentaose

Claim 7; Page 14-16; 23pp; Japanese.

Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase
variants (AAY17904-Y17909). The variants are derived from the
Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.
et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants
have a replacement of the 57th or the 130th amino acid residue in the
sequence as compared to the wild-type. The sequences are deposited under
the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The
variants can be used to transform plants to produce maltopentaose. The
enzymes prepared from the alpha-amylase variants produce maltopentaose
in a high yield.

XX SQ Sequence 1981 BP; 341 A; 677 C; 680 G; 283 T; 0 other;

Query Match 11.5%; Score 39.6; DB 20; Length 1981;

Best Local Similarity 52.4%; Pred. No. 0.22;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GCAGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62

Db 792 GCAGGTACAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851

QY 63 GCTGAGGTTGGCTTCTCCAGACTCCGCTCAGCGCAAGTCCAGTCCAAATGGCAA 122

Db 852 CGGCAAGCTGGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911

QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAAGGTCAAGGTCAAGGTCAAGGTCAAGGT 168

Db 912 GGCCATCGCTTTCGTGCAACACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 957

Search completed: June 14, 2003, 23:43:01

Job time : 80.43 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	39.6	11.5	1981	3	US-09-017-706-3	Sequence 3, Appli
2	39.6	11.5	1981	3	US-09-017-706-4	Sequence 4, Appli
3	39.6	11.5	1981	3	US-09-017-706-5	Sequence 5, Appli
4	39.6	11.5	1981	3	US-09-017-706-6	Sequence 6, Appli
5	39.6	11.5	1981	3	US-09-017-706-7	Sequence 7, Appli
6	39.6	11.5	1981	3	US-09-017-706-8	Sequence 8, Appli
7	36.4	10.6	2267	4	US-08-679-645-25	Sequence 25, Appli
8	36.4	10.6	4800	3	US-08-941-445A-4	Sequence 4, Appli
9	34.6	10.0	5267	3	US-08-976-255-2	Sequence 2, Appli
10	34	9.9	13842	4	US-09-105-537-30	Sequence 30, Appli
11	34	9.9	36778	4	US-09-105-537-5	Sequence 5, Appli
12	34	9.9	38506	3	US-09-320-878-19	Sequence 19, Appli
13	34	9.9	43280	2	US-08-804-227C-1	Sequence 1, Appli
14	33	9.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
15	33	9.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
16	32.8	9.5	1590	4	US-09-504-358-21	Sequence 21, Appli
17	32.8	9.5	1590	4	US-09-954-314-21	Sequence 21, Appli
18	32.8	9.5	11471	4	US-09-504-358-16	Sequence 16, Appli
19	32.8	9.5	11471	4	US-09-954-314-16	Sequence 16, Appli
20	32.4	9.4	1293	2	US-08-924-440-1	Sequence 1, Appli
21	32	9.3	3624	1	US-07-951-715A-6	Sequence 6, Appli
22	32	9.3	3624	2	US-08-459-448A-6	Sequence 6, Appli
23	32	9.3	3624	3	US-08-459-595A-6	Sequence 6, Appli
24	32	9.3	3624	3	US-08-459-504B-6	Sequence 6, Appli
25	32	9.3	3624	3	US-08-459-544-6	Sequence 6, Appli
26	32	9.3	3624	3	US-09-053-549-7	Sequence 7, Appli
27	32	9.3	3624	4	US-09-547-422-6	Sequence 6, Appli

RESULT 2
US-09-017-706-4
; Sequence 4, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOBENTROSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; OTHER INFORMATION: PLASMID POS3410H57
US-09-017-706-4

Query Match 11.5%; Score 39.6; DB 3; Length 1981;
Best Local Similarity 52.4%; Pred. No. 0.017;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
DB 792 GCAGGTACGGTGACCGAGTTCCGCTACGGCAAGGAGCTCTACGGCAAGTTCCGCGGGG 851
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCGCTACGGCAAGTCCAGCTCCAAAATGGCAA 122
DB 852 CGCGAAGCTGGCGACCTCGACACTTCGCGCCCGAGCTGGAACCTGATGCCCGCAGCAA 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCCGGGGC 168
DB 912 GGCCATCGCTTTCGTGCAACACCAAGCAGCAGCGCGCCACGGC 957

RESULT 3
US-09-017-706-5
; Sequence 5, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOBENTROSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)

; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID POS3410OL57
US-09-017-706-5

Query Match 11.5%; Score 39.6; DB 3; Length 1981;
Best Local Similarity 52.4%; Pred. No. 0.017;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
DB 792 GCAGGTACGGTGACCGAGTTCCGCTACGGCAAGGAGCTCTACGGCAAGTTCCGCGGGG 851
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCGCTACGGCAAGTCCAGCTCCAAAATGGCAA 122
DB 852 CGCGAAGCTGGCGACCTCGACACTTCGCGCCCGAGCTGGAACCTGATGCCCGCAGCAA 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCCGGGGC 168
DB 912 GGCCATCGCTTTCGTGCAACACCAAGCAGCAGCGCGCCACGGC 957

RESULT 4
US-09-017-706-6
; Sequence 6, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOBENTROSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID: POS3410F139
US-09-017-706-6

Query Match 11.5%; Score 39.6; DB 3; Length 1981;
Best Local Similarity 52.4%; Pred. No. 0.017;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
DB 792 GCAGGTACGGTGACCGAGTTCCGCTACGGCAAGGAGCTCTACGGCAAGTTCCGCGGGG 851
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCGCTACGGCAAGTCCAGCTCCAAAATGGCAA 122
DB 852 CGCGAAGCTGGCGACCTGCGACACTTCGCGCCCGAGCTGGAACCTGATGCCCGCAGCAA 911
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCCGGGGC 168


```
; REFERENCE/DOCKET NUMBER: 219/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-679-645-25

Query Match          10.6%; Score 36.4; DB 4; Length 2267;
Best Local Similarity 51.2%; Pred. No. 0.15;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 3 CGAGTACAAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGGCT 62
DB 1094 GAAGCGCGGATCTTCGAGCGCGACAGGTCCTCACCGTCAGCCCTACTACGCGGAGA 1153
QY 63 GCTGAGGTTTGCTTCTCCAGGATCCGGTCAAGCAAGTCCAGCTCCAAAATGGCAA 122
DB 1154 GCTCATCTCCGCGATCGCGGGGCTGCGAGCTCGACAAATCATGCGGCTCACCGGCAT 1213
QY 123 GAGCGAAGACTTCTTACATCAAGTCAAGTCAAGCAAGCCCGGGG 168
DB 1214 CACCGGCATCGTCAACGGCATGCGTCAAGGAGTGGGACCCAGC 1259

RESULT 8
US-08-941-445A-4
; Sequence 4, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(1449..1553, 1685..1765, 1860..1958, 2055
; LOCATION: ..2144, 2226..2289, 2413..2513, 2651..2760, 2858
; LOCATION: ..3101, 3212..3394, 3490..3681, 3793..3879, 3977
; LOCATION: ..4105, 4227..4343)
US-08-941-445A-4

Query Match          10.6%; Score 36.4; DB 3; Length 4800;
Best Local Similarity 51.2%; Pred. No. 0.2;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 3 CGAGTACAAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGGCT 62
DB 2897 GAAGCGCGGATCTTCGAGCGCGACAGGTCCTCACCGTCAGCCCTACTACGCGGAGA 2956
QY 63 GCTGAGGTTTGCTTCTCCAGGATCCGGTCAAGCAAGTCCAGCTCCAAAATGGCAA 122
DB 2957 GCTCATCTCCGCGATCGCGGGGCTGCGAGCTCGACAAATCATGCGGCTCACCGGCAT 3016
QY 123 GAGCGAAGACTTCTTACATCAAGTCAAGTCAAGCAAGCCCGGGG 168
DB 3017 CACCGGCATCGTCAACGGCATGCGTCAAGGAGTGGGACCCAGC 3062

RESULT 9
US-08-976-255-2/c
; Sequence 2, Application US/08976255
; Patent No. 6136581
; GENERAL INFORMATION:
; APPLICANT: Jono, Keith E.
; APPLICANT: Florman, Gregory
; TITLE OF INVENTION: KINASE GENES AND USES
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,255
; FILING DATE: No. 6136581ember 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/031,675
; FILING DATE: No. 6136581ember 22, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 229/182
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-976-255-2
```

Query Match 10.0%; Score 34.6; DB 3; Length 5267;
Best Local Similarity 52.4%; Pred. No. 0.69;
Matches 76; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 58 GGGCTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGCTCCAAATG 117
DB 2958 GAGTGGAGGCGGTGCTTCGGGCGCGGGCCCTACCTCTGAGCGCAGCTCCGCAAG 2899
QY 118 GCAAGAGCGAAGACTTCTTACATCAAGGTGAGCCAGAAAGCCGGGGCTCCCATCAC 177
DB 2898 GCCTGGGGCTCACACCCCTTCTTGGCGCTCTCTTGGAGCAGAACTCAGGGGACTCATAGTTC 2839
QY 178 CCAGATTACACGGCACTGTCCAGCG 202
DB 2838 TCGGTGTATAGCCACTGTCCAGCG 2814

RESULT 10
US-09-105-537-30

; Sequence 30, Application US/09105537A

; Patent No. 6265202

; GENERAL INFORMATION:

; APPLICANT: Sherman, D.H.

; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.

; APPLICANT: Zhao, L.

; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438US1

; CURRENT APPLICATION NUMBER: US/09/105,537A

; CURRENT FILING DATE: 1998-06-26

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 30

; LENGTH: 13842

; TYPE: DNA

; ORGANISM: Streptomyces venezuelae

US-09-105-537-30

Query Match 9.9%; Score 34; DB 4; Length 13842;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139
DB 5951 CCGAGCGCTACTGGCTCGAGAACAATCCCGCGCCCTGCGCCACCGCGCAGCACTGGCGCT 6010
QY 140 ACATCAAGTCAAGCCAGAAAGCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199
DB 6011 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG 6070
QY 200 GCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCGTCACACCC 249
DB 6071 GCGGCTGGCTCGCGCTACGCGGAGGACCACTCCGCGCAGCGCGCGCC 6120

RESULT 11
US-09-105-537-5

; Sequence 5, Application US/09105537A

; Patent No. 6265202

; GENERAL INFORMATION:

; APPLICANT: Sherman, D.H.

; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.

; APPLICANT: Zhao, L.

; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438US1

; CURRENT APPLICATION NUMBER: US/09/105,537A

; CURRENT FILING DATE: 1998-06-26

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 36778

; TYPE: DNA

RESULT 13

US-08-804-227C-1

; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 9.9%; Score 34; DB 4; Length 36778;
Best Local Similarity 50.0%; Pred. No. 2;
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139
DB 7692 CCGAGCGCTACTGGCTCGAGAACAATCCCGCGCCCTGCGCCACCGCGCAGCACTGGCGCT 7751
QY 140 ACATCAAGTCAAGCCAGAAAGCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199
DB 7752 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG 7811
QY 200 GCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCGTCACACCC 249
DB 7812 GCGGCTGGCTCGCGTCAAGCGGAGGACCACTCCGCGCAGCGCGCGCC 7861

RESULT 12

US-09-320-878-19

; Sequence 19, Application US/09320878A

; Patent No. 6117659

; GENERAL INFORMATION:

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 30062202120

; CURRENT APPLICATION NUMBER: US/09/320,878A

; CURRENT FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: CIP OF 09/141,908

; EARLIER FILING DATE: 1998-08-28

; EARLIER APPLICATION NUMBER: CIP OF 09/073,538

; EARLIER FILING DATE: 1998-05-06

; EARLIER APPLICATION NUMBER: CIP OF 08/846,247

; EARLIER FILING DATE: 1997-04-30

; EARLIER APPLICATION NUMBER: 60/119,139

; EARLIER FILING DATE: 1999-02-08

; EARLIER APPLICATION NUMBER: 60/100,880

; EARLIER FILING DATE: 1998-09-22

; EARLIER APPLICATION NUMBER: 60/087,080

; EARLIER FILING DATE: 1998-05-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 19

; LENGTH: 38506

; TYPE: DNA

; ORGANISM: Streptomyces venezuelae

US-09-320-878-19

Query Match 9.9%; Score 34; DB 3; Length 38506;
Best Local Similarity 50.0%; Pred. No. 2.1;
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139
DB 5834 CCGAGCGCTACTGGCTCGAGAACAATCCCGCGCCCTGCGCCACCGCGCAGCACTGGCGCT 5893
QY 140 ACATCAAGTCAAGCCAGAAAGCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199
DB 5894 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG 5953
QY 200 GCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCGTCACACCC 249
DB 5954 GCGGCTGGCTCGCGTCAAGCGCGGAGGAGCACTCCGCGCAGCGCGCGCC 6003

; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14351..19945
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20010..31199
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31232..36067
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36249..41774
; US-08-804-227C-1

Query Match 9.9%; Score 34; DB 2; Length 43280;
Best Local Similarity 59.2%; Pred. No. 2.1;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 213 GGGCCAGGTGGGGTGGACTTTGACCGGTCCACACCCCGGAGCTCATGCCCTTCTCCAA 272
DB 15004 GGCACGTGGCGGGGGGACCGTGTCTCACCCCGGCGTCTCGTGGACTTCTCCG 15063
QY 273 TCAGCTAGAAATGGTAAAGCGGGGGTGGTGGCAAGGT 310
DB 15064 CCAGCGCGTCTGGCCCCGACGCGGCTGCAAGGCT 15101

RESULT 14
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Query Match 9.6%; Score 33; DB 4; Length 4403765;
Best Local Similarity 49.7%; Pred. No. 18;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 75 CTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGCAAGCGGAGACTT 134
DB 3167807 CCTCAACGAGGTCTATCTGCCATGGAATCCCGACTCGACGGTGTATCCGACGGGACAT 3167748
QY 135 CTTCTACATCAAGGTACAGCCAGAAAGCCGGGGCTCCCATCACCAGATTACACGGCACT 194
DB 3167747 CGTCAACATCGACGTACCGCTACATCGTGGGGTGCACGGTGCACACCAACGCGACGT 3167688
QY 195 GTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCGTC 243
DB 3167687 TCCGGCGGGGATGTGCGACAGCAACACCGGTTGCTGTCACCGGACC 3167639

RESULT 15
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
; US-09-103-840A-1

Query Match 9.6%; Score 33; DB 4; Length 4411529;
Best Local Similarity 49.7%; Pred. No. 18;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 75 CTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGCAAGCGGAGACTT 134
DB 3173694 CCTCAACGAGGTCTATCTGCCATGGAATCCCGACTCGACGGTGTATCCGACGGGACAT 3173635
QY 135 CTTCTACATCAAGGTACGCCAGAAAGCCGGGGTCCCATCACCAGATTACACGGCACT 194
DB 3173634 CGTCAACATCGACGTACCGGCTATCATCGTGGGGTGCACGGTGCACCAACGCGACGT 3173575

Qy 195 GTCCAGCGGGATTAGGGGCCAGGCTGGGTGGACTTTGACCCGTC 243
Db 3173574 TCCGGCCGGCGATGTCGAGACGACACCGGTTGCTCGTTGACCGGACC 3173526

Search completed: June 15, 2003, 03:45:44
Job time : 36.2431 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 02:18:56 ; Search time 50.2221 Seconds
(without alignments)
9947.427 Million cell updates/sec

Title: US-09-513-888c-1_COPY_112_456
Perfect score: 345
Sequence: 1 tcgcagtagaagctgcgca.....gcaggagaaagcaaaatggg 345

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	33.3	2411	9	US-09-764-904-89
2	115	33.3	2411	9	US-10-091-548-89
3	115	33.3	2411	9	US-10-074-095-1133
4	115	33.3	2411	10	US-09-764-860-1133
5	36.4	10.6	2267	9	US-09-961-077-25
6	36	10.4	560	9	US-09-764-904-17
7	36	10.4	560	9	US-10-091-548-17
8	36	10.4	560	9	US-10-074-095-269
9	36	10.4	560	10	US-09-764-860-269
10	35	10.1	328	10	US-09-878-574-2638
11	34	9.9	13842	9	US-09-860-846-30
12	34	9.9	13842	9	US-09-988-384B-30
13	34	9.9	13842	9	US-09-836-821-30
14	34	9.9	13842	10	US-09-861-289-30
15	34	9.9	36778	9	US-09-860-846-5
16	34	9.9	36778	9	US-09-836-821-5
17	34	9.9	36778	10	US-09-861-289-5
18	34	9.9	37948	9	US-09-988-384B-5
19	34	9.9	38506	9	US-09-793-708-19

C	20	34	9.9	48841	9	US-09-844-653-32	Sequence 32, Appl
	21	33.2	9.6	2521	10	US-09-925-300-521	Sequence 521, Appl
	22	32.8	9.5	1590	10	US-09-954-314-21	Sequence 21, Appl
	23	32.8	9.5	11471	10	US-09-954-314-16	Sequence 16, Appl
	24	32.4	9.4	1883	12	US-10-033-109-9	Sequence 9, Appl
	25	32.4	9.4	2313	9	US-09-887-527-6	Sequence 6, Appl
	26	32.4	9.4	2412	9	US-10-037-270-881	Sequence 881, Appl
C	27	32.2	9.3	1985	10	US-09-864-761-15510	Sequence 15510, A
	28	32	9.3	2079	9	US-10-076-816-3	Sequence 3, Appl
	29	32	9.3	2079	9	US-10-077-381-3	Sequence 3, Appl
	30	32	9.3	3624	9	US-09-988-462-6	Sequence 6, Appl
	31	31.8	9.2	768	9	US-09-978-303-6	Sequence 6, Appl
C	32	31.8	9.2	926	9	US-10-198-846-13362	Sequence 13362, A
	33	31.4	9.1	999	9	US-09-925-664-46	Sequence 46, Appl
	34	31.4	9.1	1278	9	US-10-016-647-1	Sequence 1, Appl
	35	31.4	9.1	1844	9	US-10-016-647-3	Sequence 3, Appl
	36	31.4	9.1	3526	9	US-10-037-270-182	Sequence 182, App
	37	31.2	9.0	455	9	US-09-918-995-2482	Sequence 2482, App
	38	31.2	9.0	839	9	US-10-029-905-11	Sequence 11, Appl
	39	31.2	9.0	839	10	US-09-880-107-2223	Sequence 2223, App
	40	31	9.0	1380	10	US-09-925-301-318	Sequence 318, App
	41	31	9.0	1656	12	US-10-033-109-13	Sequence 13, Appl
	42	31	9.0	1919	9	US-10-168-066-17	Sequence 17, Appl
	43	31	9.0	2367	10	US-09-995-587A-2	Sequence 2, Appl
	44	31	9.0	2394	10	US-09-995-587A-3	Sequence 3, Appl
	45	31	9.0	2592	10	US-09-995-587A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-764-904-89
; Sequence 89, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-89

Query Match	33.3%	Score 115;	DB 9;	Length 2411;
Best Local Similarity	100.0%	Pred. No. 9e-28;		
Matches 115;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	231	CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCAATCAGCTAGAAATGGTAAG	290	
Db	1	CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCAATCAGCTAGAAATGGTAAG	60	
QY	291	CGGGGTCGCTGGCAGGTTAGTGGTTGAAACCGCAGGAAAGCAAAATGGG	345	
Db	61	CGGGGTCGCTGGCAGGTTAGTGGTTGAAACCGCAGGAAAGCAAAATGGG	115	

RESULT 2
US-10-091-548-89
; Sequence 89, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07

```
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-89

Query Match      33.3%; Score 115; DB 9; Length 2411;
Best Local Similarity 100.0%; Pred. No. 9e-28;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      231 CTTTCACCGCTCCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAAG 290
Db      1 CTTTCACCGCTCCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAAG 60

QY      291 CGGGGGTGGCTGGCAAGGTAGTGGGTTGGAAACGACGAGGAGAAAGCAAAATGGG 345
Db      61 CGGGGGTGGCTGGCAAGGTAGTGGGTTGGAAACGACGAGGAGAAAGCAAAATGGG 115

RESULT 3
US-10-074-095-1133
; Sequence 1133, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
```


;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221

;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08

Query Match 33.3%; Score 115; DB 9; Length 2411;
Best Local Similarity 100.0%; Pred. No. 9e-28;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCCAATCAGCTAGATAATGGTAAG 290
|||||
DB 1 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCCAATCAGCTAGATAATGGTAAG 60
|||||

QY 291 CGGGGGTCGCTGGCAAGGGTAAGTGGTTGGAACGAGGAGAAAGCAAAATGGG 345
|||||
DB 61 CGGGGGTCGCTGGCAAGGGTAAGTGGTTGGAACGAGGAGAAAGCAAAATGGG 115
|||||

RESULT 4
US-09-764-860-1133
; Sequence 1133, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1133
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1133

Query Match 33.3%; Score 115; DB 10; Length 2411;
Best Local Similarity 100.0%; Pred. No. 9e-28;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCCAATCAGCTAGATAATGGTAAG 290
|||||
DB 1 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTTCCAATCAGCTAGATAATGGTAAG 60
|||||

QY 291 CGGGGGTCGCTGGCAAGGGTAAGTGGTTGGAACGAGGAGAAAGCAAAATGGG 345
|||||
DB 61 CGGGGGTCGCTGGCAAGGGTAAGTGGTTGGAACGAGGAGAAAGCAAAATGGG 115
|||||

RESULT 5
US-09-961-077-25
; Sequence 25, Application US/09961077
; Publication No. US20030014775A1
; GENERAL INFORMATION:
; APPLICANT: Zwick, Michael G.
; Edington, Brent E.
; McSwiggen, James A.
; Merlo, Patricia Ann Owens
; Guo, Lining
; Skokut, Thomas A.
; Young, Scott A.
; Folkerts, Otto
; Merlo, Donald J.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; MODULATION OF GENE EXPRESSION
; IN PLANTS
; NUMBER OF SEQUENCES: 1263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/961,077
FILING DATE: 21-Sep-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/679,645
FILING DATE: July 12, 1996
APPLICATION NUMBER: 60/001,135
FILING DATE: July 13, 1995
APPLICATION NUMBER: 08/300,726
FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 219/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 2267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-961-077-25

Query Match 10.6%; Score 36.4; DB 9; Length 2267;
Best Local Similarity 51.2%; Pred. No. 0.057;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 3 CGAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62
DB 1094 GAAGCCGGGATCTTCGAGCGCGACAGGCTCTACCGTACGCCCTACTAGCGGAGA 1153
QY 63 GGTGAGTTTGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCAAAATGGCAA 122
DB 1154 GCTCATCTCCGGCATCGCCAGGGGTGCGAGCTCGACAAATCATGCGGCTCACCAGCAT 1213
QY 123 GAGCGAGACTTCTTACATCAAGTCAAGTCAAGCAAGCCCGGGC 168
DB 1214 CACCGCATCGTCAACGGCATGAGCTCAGCGAGTGGGACCCAGC 1259

RESULT 6
US-09-764-904-17
Sequence 17, Application US/09764904
Patent No. US20020173454A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA122
CURRENT APPLICATION NUMBER: US/09/764,904
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 17
LENGTH: 560
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-904-17

Query Match 10.4%; Score 36; DB 9; Length 560;
Best Local Similarity 88.6%; Pred. No. 0.053;
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 251 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTAAGCGGG 294
DB 1 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTCCGAG 44
RESULT 7
US-10-091-548-17
Sequence 17, Application US/10091548
Publication No. US20030049703A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA122C1
CURRENT APPLICATION NUMBER: US/10/091,548
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 137
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 17
LENGTH: 560
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-548-17

Query Match 10.4%; Score 36; DB 9; Length 560;
Best Local Similarity 88.6%; Pred. No. 0.053;
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 251 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTAAGCGGG 294
DB 1 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTCCGAG 44

RESULT 8
US-10-074-095-269
Sequence 269, Application US/10074095
Publication No. US2003007704A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC008C1
CURRENT APPLICATION NUMBER: US/10/074,095
CURRENT FILING DATE: 2002-02-14
Prior Application Number: 09/764,860
Prior Filing Date: 2001-01-17
Prior Application Number: 60/179,065
Prior Filing Date: 2000-01-31
Prior Application Number: 60/180,628
Prior Filing Date: 2000-02-04
Prior Application Number: 60/214,886
Prior Filing Date: 2000-06-28
Prior Application Number: 60/217,487
Prior Filing Date: 2000-07-11
Prior Application Number: 60/225,758
Prior Filing Date: 2000-08-14
Prior Application Number: 60/220,963
Prior Filing Date: 2000-07-26
Prior Application Number: 60/217,496
Prior Filing Date: 2000-07-11
Prior Application Number: 60/225,447
Prior Filing Date: 2000-08-14
Prior Application Number: 60/218,290
Prior Filing Date: 2000-07-14
Prior Application Number: 60/225,757
Prior Filing Date: 2000-08-14
Prior Application Number: 60/226,868
Prior Filing Date: 2000-08-22
Prior Application Number: 60/216,647

; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064

```

; GENERAL INFORMATION:
;
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535

```

US-09-988-384B-30
; Sequence 30, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION: D.H.
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.					
; APPLICANT: Zhao, L.					
; FILE OF INVENTION: DNA encoding methymycin and pikromycin					
; FILE REFERENCE: 600.536US1					
; CURRENT APPLICATION NUMBER: US/09/988,384B					
; CURRENT FILING DATE: 2001-11-19					
; PRIOR APPLICATION NUMBER: PCT/US99/14398					
; PRIOR FILING DATE: 1999-06-25					
; PRIOR APPLICATION NUMBER: US/09/105,537					
; PRIOR FILING DATE: 1998-06-26					
; NUMBER OF SEQ ID NOS: 53					
; SEQ ID NO 30					
; LENGTH: 13842					
; TYPE: DNA					
; ORGANISM: Streptomyces venezuelae					
US-09-988-384B-30					
Query Match 9.9%; Score 34; DB 9; Length 13842;					
Best Local Similarity 50.0%; Pred. No. 0.56;					
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;					
QY	80	CCGAGACTCCGGTCACGCCAAGTCCAGCTCCAATAATGGCAAGCGGAAGACTTTCTTCT	139		
Db	5951	CCGAGCGCTACTGGCTCGAGAACAACCTCCGCGCGCCTGGCCACCGGCAGCACTGGCGCT	6010		
QY	140	ACATCAAGGTACGCCAGAAAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCA	199		
Db	6011	ACCAGCATGACTGGAAGCGCCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG	6070		
QY	200	GCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCGTCCACACCC	249		
Db	6071	GCGGTGGCTCGCGTCAAGCGGAGGACCACTCCGCGAGCGCGCGCC	6120		
RESULT 13					
US-09-836-821-30					
; Sequence 30, Application US/09836821					
; Publication No. US20030087405A1					
; GENERAL INFORMATION:					
; APPLICANT: Sherman, D.H.					
; APPLICANT: Liu, H.					
; APPLICANT: Xue, Y.					
; APPLICANT: Zhao, L.					
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin					
; FILE REFERENCE: 600.438US1					
; CURRENT APPLICATION NUMBER: US/09/836,821					
; CURRENT FILING DATE: 2001-04-17					
; PRIOR APPLICATION NUMBER: 09/105,537					
; PRIOR FILING DATE: 1998-06-26					
; NUMBER OF SEQ ID NOS: 43					
; SOFTWARE: FastSeq for Windows Version 3.0					
; SEQ ID NO 30					
; LENGTH: 13842					
; TYPE: DNA					
; ORGANISM: Streptomyces venezuelae					
US-09-836-821-30					
Query Match 9.9%; Score 34; DB 9; Length 13842;					
Best Local Similarity 50.0%; Pred. No. 0.56;					
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;					
QY	80	CCGAGACTCCGGTCACGCCAAGTCCAGCTCCAATAATGGCAAGCGGAAGACTTTCTTCT	139		
Db	5951	CCGAGCGCTACTGGCTCGAGAACAACCTCCGCGCGCCTGGCCACCGGCAGCACTGGCGCT	6010		
QY	140	ACATCAAGGTACGCCAGAAAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCA	199		
Db	6011	ACCAGCATGACTGGAAGCGCCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG	6070		
QY	200	GCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCGTCCACACCC	249		
Db	6071	GCGGTGGCTCGCGTCAAGCGGAGGACCACTCCGCGAGCGCGCGCC	6120		
RESULT 14					
US-09-861-289-30					
; Sequence 30, Application US/09861289					
; Patent No. US20020110897A1					
; GENERAL INFORMATION:					
; APPLICANT: Sherman, D.H.					
; APPLICANT: Liu, H.					
; APPLICANT: Xue, Y.					
; APPLICANT: Zhao, L.					
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin					
; FILE REFERENCE: 600.438US1					
; CURRENT APPLICATION NUMBER: US/09/861,289					
; CURRENT FILING DATE: 2001-05-18					
; PRIOR APPLICATION NUMBER: 09/105,537					
; PRIOR FILING DATE: 1998-06-26					
; NUMBER OF SEQ ID NOS: 43					
; SOFTWARE: FastSeq for Windows Version 3.0					
; SEQ ID NO 30					
; LENGTH: 13842					
; TYPE: DNA					
; ORGANISM: Streptomyces venezuelae					
US-09-861-289-30					
Query Match 9.9%; Score 34; DB 10; Length 13842;					
Best Local Similarity 50.0%; Pred. No. 0.56;					
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;					
QY	80	CCCAGGACTCCGGTCACGCCAAGTCCAGCTCCAATAATGGCAAGCGGAAGACTTTCTTCT	139		
Db	5951	CCGAGCGCTACTGGCTCGAGAACAACCTCCCGCGCGCCTGGCCACCGGCAGCACTGGCGCT	6010		
QY	140	ACATCAAGGTACGCCAGAAAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCA	199		
Db	6011	ACCAGCATGACTGGAAGCGCCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG	6070		
QY	200	GCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCGTCCACACCC	249		
Db	6071	GCGGTGGCTCGCGTCAAGCGGAGGACCACTCCGCGAGCGCGCGCC	6120		
RESULT 15					
US-09-860-846-5					
; Sequence 5, Application US/09860846					
; Patent No. US20020164742A1					
; GENERAL INFORMATION:					
; APPLICANT: Sherman, D.H.					
; APPLICANT: Liu, H.					
; APPLICANT: Xue, Y.					
; APPLICANT: Zhao, L.					
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin					
; FILE REFERENCE: 600.438US1					
; CURRENT APPLICATION NUMBER: US/09/860,846					
; CURRENT FILING DATE: 2001-05-18					
; PRIOR APPLICATION NUMBER: 09/105,537					
; PRIOR FILING DATE: 1998-06-26					
; NUMBER OF SEQ ID NOS: 43					
; SOFTWARE: FastSeq for Windows Version 3.0					
; SEQ ID NO 5					
; LENGTH: 36778					
; TYPE: DNA					
; ORGANISM: Streptomyces venezuelae					
US-09-860-846-5					
Query Match 9.9%; Score 34; DB 9; Length 36778;					
Best Local Similarity 50.0%; Pred. No. 0.73;					
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;					
QY	80	CCCAGGACTCCGGTCACGCCAAGTCCAGCTCCAATAATGGCAAGCGGAAGACTTTCTTCT	139		
Db	7692	CCGAGCGCTACTGGCTCGAGAACAACCTCCCGCGCGCCTGGCCACCGGCAGCACTGGCGCT	7751		
QY	140	ACATCAAGGTACGCCAGAAAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCA	199		

Db 7752 ACCGATCGACTGGAAGCGCCTCCCGCGCGCCGAGGGTCCGAGCGCACCGGCTGTCCG 7811
Qy 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTGACCCGTCACACCC 249
Db 7812 GCGGCTGGCTCGCCGTCACGCGGAGGACCACTCCGCGCAGGCGCGGCC 7861

Search completed: June 15, 2003, 06:49:10
Job time : 52.2221 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 480.339 Seconds
(without alignments)

11632.279 Million cell updates/sec

Title: US-09-513-888c-1_COPY_112_456

Perfect score: 345

Sequence: 1 tcgcagtcacagctgcgcaa.....gcaggagaagcaaatggg 345

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

RST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estmu:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	232.8	67.5	658	17	AZ858490
C 2	232.4	67.4	472	12	BF706253
C 3	220	63.8	596	10	AV663618
C 4	158.8	46.0	312	17	AZ595691
C 5	122.2	35.4	817	9	A1984777
C 6	99.2	28.8	668	12	BF058214

7	62.6	18.1	72	10	BE048251
8	38.8	11.2	457	10	AV629798
9	38.4	11.1	366	13	BI064750
C 10	38	11.0	477	12	BG792715
C 11	37.6	10.9	759	14	BM675114
C 12	37	10.7	521	12	BG280616
C 13	37	10.7	535	12	BG280615
14	36.8	10.7	565	13	BJ268837
15	36.6	10.6	601	9	AL594733
16	36.6	10.6	624	9	AJ274159
17	36.6	10.6	644	9	AL783590
18	36.4	10.6	811	12	BF254153
19	36.4	10.6	2147	11	AY109531
20	35.8	10.4	540	9	AA240027
21	35.6	10.3	369	13	BJ316603
22	35.6	10.3	396	10	BE399579
23	35.6	10.3	433	9	AA485738
24	35.6	10.3	434	13	BJ213737
25	35.6	10.3	445	13	BJ222791
26	35.6	10.3	472	9	AL830730
27	35.6	10.3	476	10	BE400527
28	35.6	10.3	482	13	BM134430
29	35.6	10.3	519	9	AL830742
30	35.6	10.3	570	13	BJ289530
31	35.6	10.3	589	13	BM136270
32	35.6	10.3	600	13	BJ243353
33	35.6	10.3	604	12	BG604587
34	35.6	10.3	619	13	BJ209770
35	35.6	10.3	625	14	BQ609148
36	35.6	10.3	630	13	BJ289134
37	35.6	10.3	633	10	BE584071
38	35.6	10.3	645	14	BQ238903
39	35.6	10.3	649	13	BJ300787
40	35.6	10.3	658	13	BJ293180
41	35.6	10.3	669	13	BJ236056
42	35.6	10.3	728	13	BJ257903
43	35.6	10.3	814	9	AL820181
44	35.4	10.3	464	10	BE399162
C 45	35.4	10.3	476	13	BI340151

ALIGNMENTS

RESULT 1
AZ858490/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ858490 658 bp DNA linear GSS 21-FEB-2001
2M0163L10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0163L10 R, DNA sequence.

AZ858490
GI:13051689

GSS
house mouse.
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 658)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0163 row: L column: 10
 Seq primer: CACACAGGAACAGTATGACC
 Class: plasmid ends
 High quality sequence stop: 658.
 Location/Qualifiers
 1. 658
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0163L10"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="B. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 140 a 173 c 177 g 168 t
 ORIGIN

Query Match 67.5%; Score 232.8; DB 17; Length 658;
 Best Local Similarity 84.7%; Pred. No. 3.4e-57;
 Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
 |||||
 Db 522 TCACAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAATCGATATCTAGATGG 463
 |||||
 Qy 61 CTGCTGAGTTGGCTTCTCCAGACTCCGTCACGGCAAGTCCAGCTCCAAATGGGC 120
 |||||
 Db 462 CTGCTGAGTTGGCTTCTCCAGACTCAGGCCGTGGAAGTCAAGTTCCAAATGGGA 403
 |||||
 Qy 121 AAGAGCAAGACTTCTTCTACATCAAGTTCAGCCAGAAAGCCGGGGCTCCCATCACC 180
 |||||
 Db 402 AAAGCGAAGACTTCTTCTACATCAAGTTCAGCCAGAGGCCCGAGGCTCCCATCGCC 343
 |||||
 Qy 181 GATTACAGGCACTCTCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 240
 |||||
 Db 342 GACTACACAGCCCTCTCCAGTGGGACATAGGGGTCAGACCGGAGTAGATTTTGATCCA 283
 |||||
 Qy 241 TCCACACCCCAAGCTATGCCCTTCTCCATCAGCTAGAAATGGTAAGGGGGTGGC 300
 |||||
 Db 282 GCCACCCCAAGAGCTATGCCCTTCTCCATCACTGGAATGGTAGTGGTTCCT 223
 |||||
 Qy 301 TGGCAAGG 308
 |||||
 Db 222 TGACAGGG 215

RESULT 2
 BF706253
 LOCUS BF706253 472 bp mRNA linear EST 25-APR-2001
 DEFINITION 280459 MARC 3BOV Bos taurus cdna 5', mRNA sequence.
 ACCESSION BF706253
 VERSION BF706253.1 GI:11997851
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C., Partee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cdna libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCATGACCAT
 BACKWARD: GTTTCACAGTCACGACG
 Plate: 77 row: A column: 4
 Seq primer: ATTAGTGACACTATAG.
 Location/Qualifiers
 1. 472
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="vector: pCMV SPORT6; Site.1: NotI; Site.2: Sall;
 Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 92 a 171 c 123 g 86 t
 ORIGIN

Query Match 67.4%; Score 232.4; DB 12; Length 472;
 Best Local Similarity 92.1%; Pred. No. 4e-57;
 Matches 245; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
 |||||
 Db 207 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 266
 |||||
 Qy 61 CTGCTGAGTTGGCTTCTCCAGACTCCGTCACGGCAAGTCCAGTCCAAATGGGC 120
 |||||
 Db 267 CTGCTGAGTTGGCTTCTCCAGACTCGGGAATGCGAAGTCCAGTCCAAATGGGC 326
 |||||
 Qy 121 AAGAGCAAGACTTCTTCTACATCAAGTTCAGCCAGAAAGCCGGGGCTCCCATCACC 180
 |||||
 Db 327 AAGAGCAAGACTTCTTCTACATCAAGTTCAGCCAGAAAGCCGGGGCTCCCATCG 386
 |||||
 Qy 181 GATTACAGGCACTCTCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 240
 |||||
 Db 387 GATTACAGGCACTCTCAGTGGGGACCTAGGGGGCCAGGAGGAGTGGACTTTGACCCA 446
 |||||
 Qy 241 TCCACACCCCAAGCTCATGCCCTT 266
 |||||
 Db 447 TCCACACCCCAAGCTCATGCCCTT 472

RESULT 3
 AV663618
 LOCUS AV663618 596 bp mRNA linear EST 28-NOV-2001
 DEFINITION AV663618 Bos taurus brain fetus Bos taurus cdna clone E1BR037F07
 5', mRNA sequence.
 ACCESSION AV663618
 VERSION AV663618.1 GI:9922648


```

KEYWORDS EST.
SOURCE Bos taurus
ORGANISM cow.

REFERENCE
AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jithozono,A., Suzuki,H., Aso,H.
and Sugimoto,Y.
TITLE Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570354
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cococ.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES
source
1..596
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1BR037F07"
/clone_lib="Bos taurus brain fetus"
/tissue_type="brain"
/dev_host="fetus"
/lab_hosts="DH10B"
/note="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT 114 a 196 c 170 g 103 t 13 others
ORIGIN
Query Match 63.8%; Score 220; DB 10; Length 596;
Best Local Similarity 91.0%; Pred. No. 1.8e-53;
Matches 232; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 32 TCAGAGAGCTCAACCGGTATTCCGACGGCTGCTGAGTTGGCTTCTCCAGAGACTCG 91
Db 61 TCAAGAGCTCAACCGGTATTCCGACGGCTGCTGAGTTGGCTTCTCCAGAGACTCG 120
Qy 92 GTACGGCAAGTCCAGCTCCAAATGGCAAGAGCAAGACTTCTTACATCAAGGTCA 151
Db 121 GACATGGCAAGTCCAGCTCCAAATGGCAAGAGCAAGACTTCTTACATCAAGGTCA 180
Qy 152 GCCAGAAAGCCGGGCTCCCATCACCCAGATTACCGGCACTGTCCAGCGGGATTAG 211
Db 181 GCCAGAGGCCGGGGCTCCCAACCGCTGATTACAGGCACTGTCCAGTGGGNNNTAG 240
Qy 212 GGGGCCAGGCTGGGTGAGCTTTGACCGTTCACACCCCAAGCTATGCCCTTCTCCA 271
Db 241 GGGGCCAGGCGAGGTGAGCTTTGACCATCCACCCCAAGCTATGCCCTTCTCCA 300
Qy 272 ATCAGCTAGAAATGG 286
Db 301 ATCAGCTAGATGG 315

RESULT 4
AZ595691/c
LOCUS 312 bp DNA linear GSS 13-DEC-2000
DEFINITION IM0408A04R Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0408A04 R, DNA sequence.
ACCESSION AZ595691
VERSION AZ595691.1 GI:11717881
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 312)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0408 row: A column: 04
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 312.
Location/Qualifiers
1..312
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0408A04"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 [gi|4732114|gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 65 a 71 c 101 g 75 t
ORIGIN
Query Match 46.0%; Score 158.8; DB 17; Length 312;
Best Local Similarity 88.7%; Pred. No. 1e-35;
Matches 172; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60
Db 194 TCACAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAATCGATATCTAGATGG 135
Qy 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCCAGGCTCCAGTCCAAATGGGC 120
Db 134 CTGCTGAGGTTTGGCTTCTCCAGGACTCAGGCGGTGGAAGTCAAGTCCAAATGGGA 75
Qy 121 AAGCGGAAGACTTCTTCTATCATCAAGTCAAGCAGAAAGCCGGGGCTCCATCACC 180
Db 74 AAAAGCGAAGACTTCTTCTATCATCAAGTCAAGCAGAAAGCCGGGGCTCCATCACC 15
Qy 181 GATTACACGGCACT 194
Db 14 GACTACACAGCCCT 1

```

```
RESULT 5
AI984777/c
LOCUS
DEFINITION
  AI984777.1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:2494453 3'
  similar to TR:O60295 O60299 KIAA0552 PROTEIN. ; mRNA sequence.
ACCESSION
AI984777
VERSION
AI984777.1 GI:5812054
KEYWORDS
EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 817)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapsb@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  DNA Sequencing by: Greg Lennon, Ph.D.
  NCI Library Arrayed by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1307 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 428.
FEATURES
  source
  1..817
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:2494453"
  /clone_lib="NCI CGAP_Kid1"
  /lab_host="DH10B"
  /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
  a modified polylinker; Site 1: Not I; Site 2: Eco RI;
  Plasmid DNA from the normalized library NCI CGAP Kid3 was
  prepared, and ss circles were made in vitro. Following HAP
  purification, this DNA was used as tracer in a subtractive
  hybridization reaction. The driver was PCR-amplified cDNAs
  from a pool of 5,000 clones made from the same library
  (cloneIDs 132376-132391, 1456007-1456775, and
  1500552-1502855). Subtraction by Bento Soares and M.
  Fatima Bonaldo."
  BASE COUNT 133 a 235 c 266 g 180 t 3 others
  ORIGIN
  Query Match 35.4%; Score 122.2; DB 9; Length 817;
  Best Local Similarity 86.4%; Pred. No. 7.3e-25;
  Matches 204; Conservative 0; Mismatches 24; Indels 8; Gaps 6;
  QY 52 TCCGACGGCTGCTGAGTTGGCTTCTCCAGGACTCCGCTCAGCGCAAGTCCAGCTCC 111
  DB 756 TCCGACGGGGTGGTAAGGT--GCTTTTCCGGACTCC-GTCACGGCAAGTCAGCTCCC 700
  QY 112 AAAATGGGCAAGAGGGAAGACTTCTTACATCAAGGTCAGCCAGAAAGCCGGGGCTCC 171
  DB 699 AAA--TGGCAAGAGCGAGACTTCTTACATCAAGG-CAGCCAGAAAGCCGGGGCTCC 643
  QY 172 CATCACCGAGTACACGCACTGTCCAGCGGGGATTTAGGGGGCCAGGCT-GGGGTGGA 230
  DB 642 CATCACCGAGTAACAC-GCAGCGTCCAGCGGGGATTTAGGGGGCCAGGTTGGGGGTGAA 584
  QY 231 CTTTGACCGTCCACACCCCAAGCTCATGCTTCTCCATCAGCTAGAAATGG 286
  DB 583 CTNTGCCCGTCCACCCCAAGGCTCATGCTTCTCCATCAGCTAGAAATGG 528
RESULT 6
```

```
BF058214/c
LOCUS
DEFINITION
  BF058214.1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:3476769 3'
  similar to TR:Q9Y5W1 Q9Y5W1 FEZ1. [5] TR:Q9Y5W0 TR:Q9Y5V9 TR:Q9Y5V8
  TR:Q9Y5V7 ; contains MER22.t3 TAR1 repetitive element ; mRNA
  sequence.
ACCESSION
BF058214
VERSION
BF058214.1 GI:10812110
KEYWORDS
EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 668)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapsb@mail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
  R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
  Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
  I.M.A.G.E. Consortium DNA Sequencing by: Washington University
  Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL, send email to:
  info@image.llnl.gov
  Seq primer: -40UP from Gibco
  High quality sequence stop: 441.
FEATURES
  source
  1..668
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:3476769"
  /clone_lib="NCI CGAP Ov18"
  /tissue_type="fibrotheoma"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: ovary; Vector: pT73D-Pac (Pharmacia) with a
  modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
  strand cDNA was primed with a Not I - oligo(dT) primer [5';
  TGTTCACATCTGAAGTGGAGCGCGCGCGACATTTTTTTTTTTT 3'];
  double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not
  I and Eco RI sites of the modified pT73 vector. Library
  went through one round of normalization, and was
  constructed by Bento Soares and M. Fatima Bonaldo."
  BASE COUNT 105 a 185 c 226 g 150 t 2 others
  ORIGIN
  Query Match 28.8%; Score 99.2; DB 12; Length 668;
  Best Local Similarity 78.8%; Pred. No. 3.5e-18;
  Matches 130; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
  QY 130 GACTTCTTCTACATCAAGGTCCAGCAAGACCCGGGGTCCCATCCACCCAGATTACACG 189
  DB 668 GACTTCTTACATCATGTAGCCAGTAAAGCCGAGGCTCCATCCCAAGATTATCG 609
  QY 190 GCATGTCCAGCGGGGATTTAGGGGGCCAGCTGGGTGGAGCTTTGACCCCTCCACACC 249
  DB 608 GCATGTCCAGCGGAAGTAAAGAGGGCCAGATGGGTGAATATTGACCNCTCCACACCA 549
  QY 250 CCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGTAAGCGG 294
  DB 548 CCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGTAAGCGG 505
RESULT 7
BE048251
LOCUS
DEFINITION
  BE048251.1 NCI CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2291902 5',
  mRNA sequence.
ACCESSION
BE048251
```

VERSION KEYWORDS SOURCE	BE048251.1 EST. human.	GI:8365304
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 72)	
AUTHORS	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index	
JOURNAL	Unpublished (1998)	
COMMENT	Other ESTs: tz49a12.x1 Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Insert Length: 218 Std Error: 0.00 Seq primer: -40RP from Gibco.	
FEATURES	Location/Qualifiers	
source	1..72 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2291902" /clone_libs="NCI-CGAP_Brn52" /tissue_type="tumor, 5 pooled (see description)" /lab_host="DH10B" /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; This library represents the normalized version of NCI CGAP_Brn52. Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.19 kb. Tumor types include: meningioma, oligodendroglioma, astrocytoma (grade II), medulloblastoma, astrocytoma (grade IV). Constructed by Life Technologies."	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		
Query Match	18.1%; Score 62.6; DB 10; Length 72;	
Best Local Similarity	94.2%; Pred. NO. 7.1e-08;	
Matches	65; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	210 AGGGGGCAGCTGGGGTGGACTTTCACCGTCCACACCCCGTCCAGCTCATGCCCTTCTC 269	
Db	1 ACGGTCGGGGCTGGGGTGGAGCTTTGACCCGTCACACCCCGTCCAGCTCATGCCCTTCTC 60	
QY	270 CAATCAGCT 278	
Db	61 CAATCAGCT 69	
BASE COUNT	12 a 28 c 16 g 16 t	
ORIGIN		

```

source
1. 366
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pgfin.pk002.g2"
/clone_lib="normalized chicken fat cDNA library"
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMDH10B"
/note="Vector: pSPORT1"
BASE COUNT      94 a   89 c   123 g   60 t
ORIGIN

Query Match      11.1%; Score 38.4; DB 13; Length 366;
Best Local Similarity 56.2%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;
Matches 72; Conservative 0;

Qy 22 TCCTCCACCTCAAGAAGCTCAACGGTATTCCGACGGGCTGCTGAGGTTTGGCTTTCTCC 81
Db 283 TTCTCCAGATCCACAGCTGACAGAGTTCTCGATGAAGTCTGGAGTCTGGCTTTCTCC 224
Qy 82 CAGGACTCGGTCAGGCAAGTCCAGCTCCAAATGGGCAAGCAAGCAAGCTTTCTTAC 141
Db 223 CCTCACTCTTTTCTGCTCTTCTCGAGGAAGTGTGAAAGTAACAACACTTCTTCTC 164
Qy 142 ATCAAGGT 149
Db 163 CTTACGGT 156

RESULT 10
BG792715
LOCUS      BG792715          477 bp      mRNA      linear      EST 30-MAY-2001
DEFINITION UTSW H41A9-H21E10 UTSW Adult Mouse Cardiac Muscle Library Mus
            musculus cDNA clone UTSW_H41A9-H21E10, mRNA sequence.
ACCESSION  BG792715
VERSION    BG792715.1 GI:14128285
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 477)
AUTHORS    Gallardo,T.D., Schageman,J.V., Pertsemilidis,A., Garner,H.R.,
            Williams,R.S. and Shohet,R.V.
TITLE      UT Southwestern Medical Center, Adult Mouse Cardiac cDNA Library
JOURNAL    Unpublished (2001)
COMMENT    Contact: Schageman JJ
            Shohet/Garner Labs
            University of Texas Southwestern Medical Center
            6000 Harry Hines Blvd., NA2.226, Dallas, TX 75390, USA
            Tel: 214 648 1674
            Email: Jeff.Schageman@UTSouthwestern.edu
            cDNA library constructed by UTSW as a component of the Program for
            Genomic Applications (PGA) and the Reynolds Heart Disease
            Prevention grants for use in cDNA microarray experiments. Sequence
            Quality: Sequence ends were trimmed based on percentage of ambigu
            us base calls or 'N's in windowed segments. Sequencing: First-pass
            sequencing; ABI Prism 377 sequencer and analysis software.
            Seq primer: M13/pUC Reverse.
            Location/Qualifiers
                1..477
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="UTSW H41A9-H21E10"
                /clone_lib="UTSW Adult Mouse Cardiac Muscle Library"
                /sex="Pooled"
                /tissue_type="Cardiac muscle"
                /dev_stage="2 months"
                /lab_host="DH5a"

FEATURES             Source
1. 477
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="UTSW H41A9-H21E10"
/clone_lib="UTSW Adult Mouse Cardiac Muscle Library"
/sex="Pooled"
/tissue_type="Cardiac muscle"
/dev_stage="2 months"
/lab_host="DH5a"

Insertion site: TACGTCCACTGAATCTGAGTG--->. Other
information regarding entire library may be found at
http://pga.swmed.edu/Data/libraries/microarray_cdna_librar
ies.htm."
BASE COUNT      97 a   124 c   119 g   130 t   7 others
ORIGIN

Query Match      11.0%; Score 38; DB 12; Length 477;
Best Local Similarity 48.7%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 95; Conservative 0;

Qy 22 TCCTCCACCTCAAGAAGCTCAACGGTATTCCGACGGGCTGCTGAGGTTTGGCTTTCTCC 81
Db 221 TCCTGGCTCTCCAGAGGAGGAGCTGATTGATCTGCTGGAGTATGTGCCCGAG 280
Qy 82 CAGGACTCGGTCACGGCAAGTCCAGCTCCAAATGGGCAAGCAAGCAAGCTTTCTTAC 141
Db 281 ACGGTGTACCGAGTGGCCGCCACTTCACCAAGCCCAAGCTGATCACCCCTTTCATCTAC 340
Qy 142 ATCAAGGTCAAGCAAGAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCAGC 201
Db 341 ATCAAGGTANGCANGCANGCTGCTGGAGCAACTGTCTGGGGTTCCTTCACTTCCCAAG 400
Qy 202 GGGGATTTAGGGGCG 216
Db 401 GTGCTCTTGGGNGC 415

RESULT 11
BG792715
LOCUS      BG792715          759 bp      mRNA      linear      EST 27-FEB-2002
DEFINITION UI-E-EJ0-ahn-h-12-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
            UI-E-EJ0-ahn-h-12-0-UI 3', mRNA sequence.
ACCESSION  BG792715
VERSION    BG792715.1 GI:18985012
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 759)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
COMMENT    Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
            Location/Qualifiers
                1..759
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="UI-E-EJ0-ahn-h-12-0-UI"
                /clone_lib="UI-E-EJ0"
                /tissue_type="fetal eyes, lens, eye anterior segment,
                optic nerve, retina, Retina Foveal and Macular, RPE and
                Choroid"
                /dev_stage="fetal and adult"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

```

modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGRATCAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAI; optic nerve, CCATTAGAGG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG LIB=UI-E-EJ0
TAG_TISSUE=human lens
TAG_SEQ=CGATTAGCGA
BASE COUNT 173 a 156 c 183 g 246 t 1 others
ORIGIN

Query Match 10.9%; Score 37.6; DB 14; Length 759;
Best Local Similarity 55.3%; Pred. No. 3.5;
Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 199 AGCGGGGATTAGGGGGCAGGCTGGGCTGGACTTTGACCCGTCACACACCCCAAGCTC 258
Db 704 AGGCACATCGGGACCCACACTCTGCTTCAGTTGCATCATCTCCACCCCAATTC 645
Qy 259 ATGCCCTTCTCAATCAGCTGAATGTGTAAGCGGGGTGCTGCGCAAGGGTAAAGTGGGT 318
Db 644 AACTCTGACCAATACAAAGATTTTAAACAGGATTTCTTTCGAGAAAGCTGACT 585
Qy 319 TGGAAACGAGG 330
Db 584 TGGAAACGCGG 573

RESULT 12
BG280616
LOCUS 521 bp mRNA linear EST 21-FEB-2001
DEFINITION c5blinp.r1 Neurospora crassa sexual cDNA library, Uni-zap vector
system Neurospora crassa cDNA clone c5blnp 5', mRNA sequence.
ACCESSION BG280616
VERSION BG280616.1 GI:13079199
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 521)
AUTHORS Kupfer,D., Lai,H., Nelson,M. and Roe,B.
TITLE ESTs from a Neurospora crassa Sexual cDNA Library
JOURNAL Unpublished (2001)
COMMENT Other ESTs: c5blinp.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

FEATURES
source
1..521
/organism="Neurospora crassa"
/strain="wild type"
/db_xref="taxon:5141"

BASE COUNT 118 a 130 c 143 g 144 t
ORIGIN

Query Match 10.7%; Score 37; DB 12; Length 535;
Best Local Similarity 58.7%; Pred. No. 4.6;

/clone="c5blinp"
/clone.lib="Neurospora crassa sexual cDNA library, Uni-zap
vector system"
/tissue_type="perithecia (fruiting bodies)"
/dev_stages="sexual"
/lab_host="E. coli strain SOLR"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5', end of cDNA cloned into EcoRI site of pBluescript
; 3', end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT 130 a 154 c 115 g 122 t
ORIGIN

Query Match 10.7%; Score 37; DB 12; Length 521;
Best Local Similarity 58.7%; Pred. No. 4.6;
Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 30 CCTCAAGAGCTCAACGGTATTCGACGGGCTGCTGAGGTTCCTCCAGGACTC 89
Db 215 CATCATGTCGAGTAACAGGAATTACGATTTCTCTGATCAAGCTGCTTGTGATCGGAGACTC 274
Qy 90 CGGTCAACGGCAAGTCCAGCTCCAAAATGGGCAAGAGCAAGACTTCTTC 138
Db 275 GGGGTTGGCAAGTCTCTGCTGTTGCGCTTTAGCGAAGACTCATTC 323

RESULT 13
BG280615/c
LOCUS 535 bp mRNA linear EST 21-FEB-2001
DEFINITION c5blinp.fi Neurospora crassa sexual cDNA library, Uni-zap vector
system Neurospora crassa cDNA clone c5blnp 3', mRNA sequence.
ACCESSION BG280615
VERSION BG280615.1 GI:13079197
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 535)
AUTHORS Kupfer,D., Lai,H., Nelson,M. and Roe,B.
TITLE ESTs from a Neurospora crassa Sexual cDNA Library
JOURNAL Unpublished (2001)
COMMENT Other ESTs: c5blinp.r1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

FEATURES
source
1..535
/organism="Neurospora crassa"
/strain="wild type"
/db_xref="taxon:5141"
/clone="c5blinp"
/clone.lib="Neurospora crassa sexual cDNA library, Uni-zap
vector system"
/tissue_type="perithecia (fruiting bodies)"
/dev_stages="sexual"
/lab_host="E. coli strain SOLR"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5', end of cDNA cloned into EcoRI site of pBluescript
; 3', end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT 118 a 130 c 143 g 144 t
ORIGIN

Query Match 10.7%; Score 37; DB 12; Length 535;
Best Local Similarity 58.7%; Pred. No. 4.6;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 2092.33 Seconds
(without alignments)
11183.052 Million cell updates/sec

Title: US-09-513-888C-1_COPY_1707_2510
Perfect score: 804
Sequence: 1 gccatctgactctcccc.....tttggcgccagatccccct 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	804	100.0	9108	9	AF123653	AF123653 Homo sapi
2	802.4	99.8	173264	2	AC025853	AC025853 Homo sapi
3	746.2	92.8	1515	9	AF123656	AF123656 Homo sapi
4	746.2	92.8	1614	9	AF123655	AF123655 Homo sapi
5	746.2	92.8	1692	9	AF123657	AF123657 Homo sapi
6	746.2	92.8	1722	9	AF123658	AF123658 Homo sapi
7	746.2	92.8	5492	9	AF123659	AF123659 Homo sapi
8	542.2	67.4	227884	2	AC114995	AC114995 Mus muscu
9	542.2	67.4	263546	2	AC099416	AC099416 Mus muscu
10	233.2	29.0	191210	2	AC108987	AC108987 Rattus no
11	157.6	19.6	633	9	AF123654	AF123654 Homo sapi
12	152	18.9	125856	9	HS1187M17	AL121891 Human DNA
13	151	18.8	5257	9	AB011124	AB011124 Homo sapi
14	148.2	18.4	205581	10	AL731707	AL731707 Mouse DNA
15	148.2	18.4	220355	2	AL772162	AL772162 Mus muscu
16	136.6	17.0	2416	10	RNO278801	AC129762 Rattus no
17	130	16.2	159724	2	AC129762	AY029201 Homo sapi
18	125.4	15.6	1935	9	AY029201	AB046013 Macaca fa
19	125.4	15.6	2099	9	AB046013	AL834338 Homo sapi
20	125.4	15.6	2275	9	HSM805394	BC006212 Homo sapi
21	125.4	15.6	2766	9	BC006212	AB058716 Homo sapi
22	125.4	15.6	5733	9	AB058716	BC014695 Mus muscu
23	113.8	14.2	2767	10	BC014695	AX400518 Sequence
24	112.4	14.0	561	6	AX400518	AL133215 Human DNA
25	94	11.7	181086	9	AL133215	AK097997 Homo sapi
26	92.4	11.5	2308	9	AK097997	BC005855 Homo sapi
27	87.8	10.9	1612	9	BC005855	AC106674 Rattus no
28	86.4	10.7	180418	2	AC106674	AL833803 Mus muscu
29	85.4	10.6	225045	2	AL833803	AC099443 Rattus no
30	84.8	10.5	177035	2	AC099443	AC122329 Mus muscu
31	84.6	10.5	167473	2	AC122329	AC116700 Mus muscu
32	84.6	10.5	168210	2	AC116700	AC119801 Mus muscu
33	84.6	10.5	251958	2	AC119801	AC121787 Mus muscu
34	84.6	10.5	325808	2	AC121787	AC130883 Rattus no
35	84.4	10.5	157147	2	AC130883	Continuation (4 of
36	83.8	10.4	103729	2	AC096620	AC105485 Rattus no
37	83.8	10.4	270700	2	AC105485	AL606742 Mouse DNA
38	83.2	10.3	186559	10	AL606742	AC090533 Mus muscu
39	83	10.3	272545	2	AC090533	J02079 Epstein-Bar
40	82.8	10.3	1150	14	HS4ULIR3	AX107940 Sequence
41	82.8	10.3	1926	6	AX107940	AR108994 Sequence
42	82.8	10.3	2580	6	AR108994	AR083151 Sequence
43	82.8	10.3	5452	6	AR083151	U02454 Cloning vec
44	82.8	10.3	5452	12	U02454	A92665 Sequence 1
45	82.8	10.3	9600	6	A92665	

ALIGNMENTS

RESULT 1
AF123653
LOCUS
DEFINITION Homo sapiens FEZ1 (FEZ1) gene, complete cds.
ACCESSION AF123653
VERSION AF123653.1 GI:4572463
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 9108)
REFERENCE
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,

Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneu, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 17, 2002 this sequence version replaced 91:21321864.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7454
Center clone name: 353_K12

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 39461: contig of 39461 bp in length
* 39462 39561: gap of 100 bp
* 39562 108347: contig of 68786 bp in length
* 108348 108447: gap of 100 bp
* 108448 173264: contig of 64817 bp in length.

Location/Qualifiers
1. 173264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"
/clone="RP11-353K12"
/clone_lib="RP11-11 Human Male BAC"
/clone_lib="RP11-11 Human Male BAC"

BASE COUNT 47355 a 39864 c 39235 g 47493 t 317 others

ORIGIN

Query Match 99.8%; Score 802.4; DB 2; Length 173264;
Best Local Similarity 99.9%; Pred. No. 3.6e-121;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 23989 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 24048

Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCTGCCTGTCTGTGGGGCTCTCAGACTCCGGC 120
Db 24049 AAGCCCAAGGAGCAGAGCTGAAGCTGCCTGTCTGTGGGGCTCTCAGACTCCGGC 24108

Qy 121 CGGAACCTCCATGTCCAGGCTGCCACACACAGCAGCAGCAGCTACCAGCTGGACCCG 180

24109 CGGAACCTCCATGTCCAGGCTGCCACACACAGCAGCAGCAGCTACCAGCTGGACCCG 24168
181 CTGCTGCACACCCGTTGGGACCCACAAAGCGTTTGGGGCTCGGCCACCAATCACCAG 240
24169 CTGCTGCACACCCGTTGGGACCCACAAAGCGTTTGGGGCTCGGCCACCAATCACCAG 24228
241 GGCATCGTCTCTCCAGGACAGCAACATGATGAGCCTTGAAGGCTCTGTCTCTTCTCCGACCGA 300
24229 GGCATCGTCTCTCCAGGACAGCAACATGATGAGCCTTGAAGGCTCTGTCTCTTCTCCGACCGA 24288
301 GGTAGCAAGCTGGGCGCACTCGAAACAGGACAGCAAGGGCCCTCGTGTGCTGCTCCGCC 360
24289 GGTAGCAAGCTGGGCGCACTCGAAACAGGACAGCAAGGGCCCTCGTGTGCTGCTCCGCC 24348
361 ATCTCCAGGACAGTGCAGCATCAGGAGCTGGAACAGAGCTCTTCCGAGGGGAGGGC 420
24349 ATCTCCAGGACAGTGCAGCATCAGGAGCTGGAACAGAGCTCTTCCGAGGGGAGGGC 24408
421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGGCTCCAGGCTTGGCCTTAC 480
24409 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGGCTCCAGGCTTGGCCTTAC 24468
481 GAGGAGCGCCCGCGCGCTGCAGGAGAGTGTGAGGGCCCGGAGCCCAAAGGGCGGCAAC 540
24469 GAGGAGCGCCCGCGCGCTGCAGGAGAGTGTGAGGGCCCGGAGCCCAAAGGGCGGCAAC 24528
541 AAGCTCAAGCAGGCTCCGAGAGAGCAGCGCGCGGAGAGGAGCTCTGCACCTGCAGGTA 600
24529 AAGCTCAAGCAGGCTCCGAGAGAGCAGCGCGCGGAGAGGAGCTCTGCACCTGCAGGTA 24588
601 CTGCAAGCTTTCAGCAGGAGAGCAGCGCAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGGAG 660
24589 CTGCAAGCTTTCAGCAGGAGAGCAGCGCAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGGAG 24648
661 CAGCACTCTGCGAGACCAAGCTCAGGTCTTACGAGGGGAGAGACAGCTTTCGCGCCC 720
24649 CAGCACTCTGCGAGACCAAGCTCAGGTCTTACGAGGGGAGAGACAGCTTTCGCGCCC 24708
721 GCGCTGAGGAGAGCCAGTGGAGGTGAGGCCACACAGGGCTCATGGTGTGGTGGTCA 780
24709 GCGCTGAGGAGAGCCAGTGGAGGTGAGGCCACACAGGGCTCATGGTGTGGTGGTCA 24768
781 GCGGTGTTGGCGCCAGTACCCCT 804
24769 GCGGTGTTGGCGCCAGTACCCCT 24792

RESULT 3
AF123656
LOCUS Homo sapiens clone D14 FE21 (FE21) mRNA, alternatively spliced,
DEFINITION complete cds.
ACCESSION AF123656
VERSION AF123656.1 GI:4572469
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE
1 (bases 1 to 1515)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C. M.
The FE21 gene at chromosome 9p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
9199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 1515)
Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C. M.
Direct Submission
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer

Institute; 233 S 10th street, Philadelphia, PA 19107, USA

FEATURES
source Location/Qualifiers
1..1515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="D14"
gene 1..1515
/gene="FEZ1"
CDS 1..1515
/gene="FEZ1"
/note="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23837.1"
/db_xref="GI:4572470"
translations="MGSVSDLSIGHFSKHCRAOVKLKRKSHLKNRYSDGLIRF
GFSDSGHGSKSSMKVSEDFYIKVSQARGSHHPDYTALSSDGLGGGAGVDPDFST
PPKLMPFNQLEMSGEKAVRPTAFKPLVRSCAILHSPESASHOLHPAPDKPKQE
ELKPICSGALSOSGRNWSMLPHTSISSYQDLPVTPVGTSRSFGSAHNITQGIV
LDQSNMMSLKSFSDGSKLGHNSKADGPSRVSPISDTDECSIQLSGKLLERGA
LQKORSFEELASLAYEERRRCRDLEPEPGKGNKLQAOKSQORAQQVLHQ
VLQOQRKRLQDELJSLMEQDLLETKLREYRKTFGPAALEETQWEVCCKSGRIS
LLKQOLEKSTEYNAKASETLGLKAOLKOTRGKLEGLELTODLEGALRTKGLELVC
ENEOQSYYAMYORNELEXALOQLAGDSAGEPLEVDLEGADIPYEDIATRI"

BASE COUNT 336 a 483 c 482 g 214 t

ORIGIN

	Query Match	Score	746.2;	DB 9;	Length	1515;	
	Best Local Similarity	99.6%;	Pred.	No. 1e-111;			
	Matches	748;	Conservative	0;	Mismatches	3;	
				Indels	0;	Gaps	0;
Qy	1	GCATCTCGACTCTCCCGGAGGTGCCAGCACCAAGTCACAGCTGCACC	CGCGCCCTCGAC				
Dd	406	GCCATCTCGACTCTCCCGGAGGTGCCAGCACCAAGTCACAGCTGCACC	CGCGCCCTCGAC				
Qy	61	AAGCCCAGGAGGAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGT	CAGACTCCGGC	120			
Dd	466	AAGCCCAGGAGGAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGT	CAGACTCCGGC	525			
Qy	121	CGGAATCTCATGTTCAGCCTGCCACACACAGACAAGAGAGAGCTACCA	AGCTTGAGCCCG	180			
Dd	526	CGGAATCTCATGTTCAGCCTGCCACACACAGACAAGAGAGAGCTACCA	AGCTTGAGCCCG	585			
Qy	181	CTGTCTACACCCGTGGGACCCACAAGAAGCGTTTGGGGGCTCCGCCCCA	CAACATCACCAG	240			
Dd	586	CTGTCTACACCCGTGGGACCCACAAGAAGCGTTTGGGGGCTCCGCCCCA	CAACATCACCAG	645			
Qy	241	GGATCGTCTCTCCAGGACAGCAAATGATGAGCCTGAAGCTCTGTCTCTT	CTCCGACGGA	300			
Dd	646	GGATCGTCTCTCCAGGACAGCAAATGATGAGCCTGAAGCTCTGTCTCTT	CTCCGACGGA	705			
Qy	301	GGTAGAAGCTGGGCCACTCGAACAGAGGACAGAAAGGGCCCTCGTGTG	TCCGCTCCCC	360			
Dd	706	GGTAGAAGCTGGGCCACTCGAACAGAGGACAGAAAGGGCCCTCGTGTG	TCCGCTCCCC	765			
Qy	361	ATCTCCACGACGAGTGCAGCATCCAGGAGCTCGAACAGAGCTGTTGGAG	AGGGAGGGC	420			
Dd	766	ATCTCCACGACGAGTGCAGCATCCAGGAGCTCGAACAGAGCTGTTGGAG	AGGGAGGGC	825			
Qy	421	GCCTCTCAGAAGCTGCAGCGAGCTTTTGAGAGAAAGAGTGTGCTCCAG	CGCTGGCTAC	480			
Dd	826	GCCTCTCAGAAGCTGCAGCGAGCTTTTGAGAGAAAGAGTGTGCTCCAG	CGCTGGCTAC	885			
Qy	481	GAGGAGCGCGCGCGCTGCAGGAGAGCTGGAGGGCCCGGAGCCCAAGCG	CGGCAAC	540			
Dd	886	GAGGAGCGCGCGCGCTGCAGGAGAGCTGGAGGGCCCGGAGCCCAAGCG	CGGCAAC	945			
Qy	541	AAGTCTCAAGCAGGCCTCGCAAGAGAGCCAGCGCGCGCAGCAGGTCTC	GCACCTCAGGTA	600			
Dd	946	AAGTCTCAAGCAGGCCTCGCAAGAGAGCCAGCGCGCGCAGCAGGTCTC	GCACCTCAGGTA	1005			

and its expression is altered in multiple human tumors

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
10097140
2 (bases 1 to 5492)
Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.
Direct Submission
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA
Location/Qualifiers
1. 5492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
1. 5492
/gene="FEZ1"
112..1902
/gene="FEZ1"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23840.1"
/db_xref="GI:4572476"
/translation="MGSVSSLSIGHSFHSHKCRASQYKLRKSHLKKLNRYSDGLLRP
GFSQDSHGKSSXMGKSEDFYFKVQKARGSHHPDYATLSSDGLGQAGVDFDPST
PKLMPFSNQIEMSEKGAFTKPVLPKRGALHSSPESASHQHPADPKPEQ
ELKPLCSGALSDSGRMSLSLTHSTSSSYQDLPVTPVGTFRFGSAHNTIGQIV
LQDSNMSLKALSDSGSKLGNKADKPCSVRSPISTDECISIOLEOKLJEREGA
LQKLOREFEELASSLAYERPRCRDELEGPFGKGNKLKQASQSQRAQVHLIQ
VLQOQEKQRQURLESMLKQDLETKLRSYREKTSFGPALBETQWEVCQKSGEIS
LLKQOLKESEVNAKASEILGLKAQKDTGKLEGLLELRTQDLEGALRTKGLSEVC
ENLQRKNEAEELREKNLLEQLELRQAALARDMGPTTFDPFALORELERLR
ENLREERQHDQSSGFOHERLVNKEKEKVIQYQKLOQSYVAMYQNRQRLKALQ
LARGDSAGEPLEVDLEGADIPYDIIATEI"
BASE COUNT 1137 a 1704 c 1565 g 1086 t

Query Match 92.8%; Score 746.2; DB 9; Length 5492;
Best Local Similarity 99.6%; Pred. No. 8.2e-112;
Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
DB 517 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 576
QY 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTCAAGCTCCGC 120
DB 577 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTCAAGCTCCGC 636
QY 121 CGGAATCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACAGCTGAGCCCG 180
DB 637 CGGAATCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACAGCTGAGCCCG 696
QY 181 CTGCTACACCCGTGGGACCCACAAAGCGTTTGGGGGCTCCGCCCAACATCACCAG 240
DB 697 CTGCTACACCCGTGGGACCCACAAAGCGTTTGGGGGCTCCGCCCAACATCACCAG 756
QY 241 GGCATGCTCTCCAGGACGACATGATGAGCTGAGGCTGTCTCTTCGACGCGA 300
DB 757 GGCATGCTCTCCAGGACGACATGATGAGCTGAGGCTGTCTCTTCGACGCGA 816
QY 301 GGTAGCAGCTGGGCGCACTCGAAAGGACGACAGGCGCCCTCGTGTGTCGGCTCCGCC 360
DB 817 GGTAGCAGCTGGGCGCACTCGAAAGGACGACAGGCGCCCTCGTGTGTCGGCTCCGCC 876
QY 361 ATCTCCAGGACGAGTGAGCATCCAGAGCTGGAAAGAGAGCTGTGGAGGAGGAGGC 420
DB 877 ATCTCCAGGACGAGTGAGCATCCAGAGCTGGAAAGAGAGCTGTGGAGGAGGAGGC 936
QY 421 GCGCTCCAGAGCTGCAGCGCAGCTTTGAGGAGGAGGCTTGCCTCCAGCTGCCTAC 480
DB 937 GCGCTCCAGAGCTGCAGCGCAGCTTTGAGGAGGAGGAGCTTGCCTCCAGCTGCCTAC 996

QY 481 GAGGAGCGCGCGCGCTGCAGGAGCAGCTGAGGCGCCCGAGCCCAAGGGCGGCAAC 540
DB 997 GAGGAGCGCGCGCGCTGCAGGAGCAGCTGAGGCGCCCGAGCCCAAGGGCGGCAAC 1056
QY 541 AAGCTCAAGCAGGCTCGCAGAAAGCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA 600
DB 1057 AAGCTCAAGCAGGCTCGCAGAAAGCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA 1116
QY 601 CTGCAAGCTTCAGCAGGAGAGCGCAGCTCGGAGGAGCTCGAGAGCTCATGAAGGAG 660
DB 1117 CTGCAAGCTTCAGCAGGAGAGCGCAGCTCGGAGGAGCTCGAGAGCTCATGAAGGAG 1176
QY 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCTCTACAGAGGAGGAGAGACAGCTTCGCGCCC 720
DB 1177 CAGGACCTGCTGGAGACCAAGCTCAGGTCTCTACAGAGGAGGAGAGACAGCTTCGCGCCC 1236
QY 721 GCGCTGGAGGAGACCCAGTGGGAGGTGAGGC 751
DB 1237 GCGCTGGAGGAGACCCAGTGGGAGGTGAGGC 1267

RESULT 8
AC114995 227884 bp DNA linear HTG 26-JUN-2002
LOCUS Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered
DEFINITION pieces.
AC114995
AC114995.3 GI:21592059
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 227884)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-171P5
Unpublished
2 (bases 1 to 227884)
Anderson, S., Barna, N., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Boukhalil, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Farro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Jones, C.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
McCarthy, M., McDonald, P., Major, J., Marquis, N., Matthews, C.,
McLean, C., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 227884)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalil, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Farro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,

Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labrocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vasiliiev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 26, 2002 this sequence version replaced gi:21535965.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23473
Center clone name: 171_P_5
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 225162 bases at least Q40
Consensus quality: 226535 bases at least Q30
Consensus quality: 227117 bases at least Q20
Insert size: 225000; agarose-fp
Insert size: 227484; sum-of-contigs
Quality coverage: 8.6 in Q20 bases; agarose-fp
Quality coverage: 8.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced as soon as it is available and by the finished sequence as preserved.
* the accession number will be preserved.
* 1 301: contig of 301 bp in length
* 302 401: gap of 100 bp
* 402 5831: contig of 5430 bp in length
* 5832 5931: gap of 100 bp
* 5932 173029: contig of 167098 bp in length
* 173030 173129: gap of 100 bp
* 173130 210061: contig of 36932 bp in length
* 210062 210161: gap of 100 bp
* 210162 227884: contig of 17723 bp in length.
Location/Qualifiers
1. .227884

FEATURES

source

misc_feature

1. .301
/note="assembly_fragment
clone end:SP6
vector side:left"
402..5831
/note="assembly_fragment"
5932..173029
/note="assembly_fragment"

misc_feature

402..5831

misc_feature

5932..173029

misc_feature 173130..210061
/note="assembly_fragment"
210162..227884
/note="assembly_fragment
clone end:T7
vector side:right"
BASE COUNT 65784 a 48829 c 49902 g 62967 t 402 others
ORIGIN

Query Match 67.4%; Score 542.2; DB 2; Length 227884;
Best Local Similarity 83.5%; Pred. No. 5.5e-79;
Matches 629; Conservative 0; Mismatches 118; Indels 6; Gaps 1;
Qy 1 GCCATCTGCACTCCTCCCGAGAGTGCACAGCCAGCTGCACCCCGCCCTCCACAC 60
Db 84432 GCCATCTGCACTCCTCCCGAGAGTGCACAGCCAGCTGCACCCCGCCCTCCACAC 84491
Qy 61 AAGCCCAAGAGCAGGAGCTGAAGCCTGGCTCTGTGGGGGGCTGTCCAGACTCCGGC 120
Db 84492 AAGCCCAAGAGCAGGAGCTGAAGCCTGGCTCTGTGGGGGGCTGTCCAGACTCCGGC 84551
Qy 121 CGGAATCTCATGTCCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCG 180
Db 84552 CGGAATCTCATGTCCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGATCCT 84611
Qy 181 CTGTGTACACCCGTTGGGACCCACAGCCGTTTTCGGGGCTCCGCCCAACATCACCCAG 240
Db 84612 CTGTGTACACCCGTTGGGACCCACAGCCGTTTTCGGGGCTCCGCCCAACATCACCA 84671
Qy 241 GGCATCTGCTCTCCAGGACAGCAACATGATAGCTCTGAAGGCTCTGTCTTCCAGCGA 300
Db 84672 GGCATCTGCTCTCCAGGACAGCAACATGATAGCTCTGAAGGCTCTGTCTTCCAGCGA 84731
Qy 301 GGTAGCAGCTGGGGCCTCGAAACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 84732 GGTAGCAGCTGGGGCCTCGAAACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGG 84791
Qy 361 ATCTCCAGGACAGTGCAGCATCCAGGAGCTGGAACAGAGAGCTGTTGGAGGAGGAGG 420
Db 84792 ATCTCCAGGACAGTGCAGCATCCAGGAGCTGGAACAGAGAGCTGTTGGAGGAGGAGG 84851
Qy 421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGGCTCCAGGCTTGCCTTAC 480
Db 84852 GCACCTACAGAGCTACAGCGCAGTTTCATGAGAAGGAGTTTGGCTCTGGCCAAACCTTT 84911
Qy 481 GAGAGGCGCCGCGCGCTGCAGGACGAGCTGGAGGCGCCCGAGGAGGAGGAGGAGGAGG 540
Db 84912 GAGAGGCGCCGCGCGCTGCAGGACGAGCTGGAGGCGCCCGAGGAGGAGGAGGAGG 84965
Qy 541 AAGCTCAAGCAGGCTCCGAGAAGAGCCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA 600
Db 84966 AAGCTCAAGCAGGCTCCGAGAAGAGCCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA 85025
Qy 601 CTCGAGCTTACAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGGAG 660
Db 85026 TTGCAGTTTCAGCAGGAGAGAAACCGCAGCTAAGGAGGAGCTTAGAGAGCTTAATGAAGGAG 85085
Qy 661 CAGGACCTGCTGGAGAGCAACCTCAGTCTTACGAGGAGGAGAGACAGCTTCCGCGCC 720
Db 85086 CAGGACCTGCTGGAGAGCAACCTCAGTCTTACGAGGAGGAGAGACAGCTTCCGCGCC 85145
Qy 721 CGCTGGAGGAGAGCCAGTGGGAGGTGAGGCA 753
Db 85146 GCACCTGGAGGAGAGCCAGTGGGAGGTGAGGCA 85178

RESULT 9

AC099416/c
LOCUS
DEFINITION Mus musculus chromosome UNK clone RP23-122M11, WORKING DRAFT
ACCESSION AC099416
VERSION AC099416.2 GI:21105058

AC099416 263546 bp DNA linear HTG 23-MAY-2002
Mus musculus chromosome UNK clone RP23-122M11, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
AC099416
AC099416.2 GI:21105058

KEYWORDS
SOURCE
ORGANISM

HTG: HTGS_PHASE1; HTGS_DRAFT.
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 263546)
McPherson, J.D. and Waterston, R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 263546)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 263546)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On May 23, 2002 this sequence version replaced gi:16924178.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0122M11
----- Summary Statistics -----
Sequencing vector: M13; 32x
Sequencing vector: plasmid; 68x
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 278310 bases at least Q40
Consensus quality: 282938 bases at least Q30
Consensus quality: 286149 bases at least Q20
Insert size: 250000; agarose-fp
Insert size: 387570; sum-of-contigs
Quality coverage: 25.74 in Q20 bases; agarose-fp
Quality coverage: 16.86 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1204: contig of 1204 bp in length
* 1205 1304: gap of unknown length
* 1305 2524: contig of 1220 bp in length
* 2525 2624: gap of unknown length
* 2625 4136: contig of 1512 bp in length
* 4137 4236: gap of unknown length
* 4237 16434: contig of 12198 bp in length
* 16435 16534: gap of unknown length
* 16535 32106: contig of 15572 bp in length
* 32107 32206: gap of unknown length
* 32207 140224: contig of 108018 bp in length
* 140225 140324: gap of unknown length
* 140325 262693: contig of 122369 bp in length
* 262694 262793: gap of unknown length
* 262794 263546: contig of 753 bp in length.

FEATURES
source

1..263546
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-122M11"
1..1204

misc_feature

/note="assembly_name:Contig127"
1305..2524
/note="assembly_name:Contig157"
2625..4136
/note="assembly_name:Contig360"
4237..16434
/note="assembly_name:Contig396"
16535..32106
/note="assembly_name:Contig397"
32207..140224
/note="assembly_name:Contig398"
140325..262693
/note="assembly_name:Contig399"
262794..263546
/note="assembly_name:Contig154"
BASE COUNT 76628 a 56149 c 55623 g 74436 t 710 others
ORIGIN
Query Match 67.4%; Score 542.2; DB 2; Length 263546;
Best Local Similarity 83.5%; Pred. No. 5.4e-79;
Matches 629; Conservative 0; Mismatches 118; Indels 6; Gaps 1;
QY 1 GCCATCTCTGCACTCTCTCCCGAGAGTCCAGCCAGCAGCTGCACCCCGCCCTCCAGAC 60
DB 250454 GCCATCTCTCCACTTTTACCTGAGAGTACCAGCCACCAACTTCATCCCATGCCCTCCAGAT 250395
QY 61 AAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGGCTGTCTGAGACTCCGCG 120
DB 250394 AAGCCCAAGGAGCAGGAGCTGAAGCAGCCTGTGCTCTGGGGGCTGTCTGAGACTCAGGC 250335
QY 121 CGGAATCTCATGTCTCCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCCG 180
DB 250334 CGGAATCTCATGTCTAGCTCTCCCGAGCATAGCAGCAGCTACCACTGGATCCT 250275
QY 181 CTGTTCACACCCGTGGGACCCACAAAGCGTTTTGGGGCTCGGCCCAACATACCCAG 240
DB 250274 CTGTTCACACCCGTGGGACCCACAAAGCGTTTTGGGGCTCGGCCCAACATACCA 250215
QY 241 GGCATCGTCTCCAGGACAGCAATGATGAGCTGTAAGGCTCTGTCTCTCCGACCGA 300
DB 250214 GGCATCATCTTCAGGACAGCAATGATGAGCTGTAAGGCTCTGTCTCTGTGATGCG 250155
QY 301 GGTAGCAAGCTGGGCCACTCGAACAAAGCAGCAGCAGCAGCTGTGTGTCTCGCTCCCCC 360
DB 250154 GGCAGCAAGCTGGCTCACCCAGGCAAGCAGCAGCTGTGTGTCTCGCTCCCCA 250095
QY 361 ATCTCCACGACGAGTGCAGCATCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 420
DB 250094 CTCTCCAGGATGATGCACCATCCAGGAGCTGGAGCAGAGCTGCTGCAGAGGAGACT 250035
QY 421 GCCCTCCAGAAAGCTGCAGCGCAGCTTTTGAGGAGAGGAGCTTGCTCCAGCTGCCCTAC 480
DB 250034 GCATACAGAAAGCTACAGCGCAGTTTCATGAGAAAGAGTTTGCTCTGGCCAAACCTTT 249975
QY 481 GAGGAGCGCCCGCGCGCTGCAGGAGCAGCTGGAGGCGCCCGAGCCCAAGCGGCAAC 540
DB 249974 GAGGAGCGCCCGCGCGCTGCAGGAGCAGCTGGAGCTGGAGCTGGAGCTGGAGCT 249921
QY 541 AAGCTCAAGCAGGCTTCGACAGAGCAGCGCGCGAGAGGCTCTGCACCTGCAGGTA 600
DB 249920 AAGCTGAAGCCACCTTCACAGAGAGCAGCGCAGCAGAGCTGCTCCAGCTGAGGTTG 249861
QY 601 CTGCAGCTTCAGCAGGAGAGCGGCAGCTCCCGCAGGAGCTCGAGAGCTCATGAAGAG 660
DB 249860 TTGCAGTTGTCAGCAGGAGAGAAACCGCAGCTAAGGAGGAGCTTAGAGAGCTTAATGAAGAG 249801
QY 661 CAGGACCTGCTGGAGAGCAAGCTCAGTCTTACGAGAGGAGAGAGCAGCTTCGCGCCC 720
DB 249800 CAGGACCTGCTGGAGAGCAAGCTCAGTCTTACGAGAGGAGAGAGCAGCTTCGCGCCC 249741
QY 721 CGCTGGAGGAGAGCCAGTGGAGGTGAGGCA 753
DB 249740 GCACCTGGAGGAGAGCCCAATGGAGGTAGACAA 249708

RESULT 10
 AC108987
 LOCUS
 DEFINITION
 AC108987
 Rattus norvegicus clone CH230-115K1, *** SEQUENCING IN PROGRESS
 ***, 78 unordered pieces.
 AC108987.3 GI:21737647
 HTG: HTGS_PHASE1.
 KEYWORDS
 Norway rat.
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 191210)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbala,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
 Homsí,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Lucier,R., Luna,R., Ma,J.,
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maneshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okuwonu,G.,
 Oreguay,N., Oviedo,R., Pace,A., Payton,B., Peary,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Washington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 191210)
 Worley,K.C.
 TITLE
 Direct Submission
 Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 191210)
 Worley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18846600.
 ----- Genome Center
 Center: Baylor College of Medicine

Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GPMW
 Center clone name: CH230-115K1
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 115719 bases at least Q40
 Consensus quality: 122885 bases at least Q30
 Consensus quality: 129025 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 78 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1041: contig of 1041 bp in length
 * 1141: gap of unknown length
 * 1142: contig of 1121 bp in length
 * 2262: gap of unknown length
 * 2362: contig of 1176 bp in length
 * 3538: gap of unknown length
 * 3639: contig of 1137 bp in length
 * 4775: gap of unknown length
 * 4875: contig of 1175 bp in length
 * 6050: gap of unknown length
 * 6150: contig of 1510 bp in length
 * 7660: gap of unknown length
 * 7761: contig of 1571 bp in length
 * 9332: gap of unknown length
 * 9432: contig of 1596 bp in length
 * 11027: gap of unknown length
 * 11127: contig of 1642 bp in length
 * 12869: gap of unknown length
 * 12770: contig of 1223 bp in length
 * 14092: gap of unknown length
 * 14192: contig of 1036 bp in length
 * 15228: gap of unknown length
 * 15328: contig of 1381 bp in length
 * 16709: gap of unknown length
 * 16809: contig of 1386 bp in length
 * 18195: gap of unknown length
 * 18295: contig of 1117 bp in length
 * 19412: gap of unknown length
 * 19512: contig of 1447 bp in length
 * 20959: gap of unknown length
 * 21059: contig of 1102 bp in length
 * 22161: gap of unknown length
 * 22261: contig of 1326 bp in length
 * 23587: gap of unknown length
 * 23687: contig of 1544 bp in length
 * 25231: gap of unknown length
 * 25331: contig of 1804 bp in length
 * 27135: gap of unknown length
 * 27235: contig of 1940 bp in length
 * 28275: gap of unknown length
 * 30846: contig of 1571 bp in length
 * 30946: gap of unknown length
 * 32090: contig of 1144 bp in length
 * 32091: gap of unknown length
 * 34039: contig of 1849 bp in length
 * 34139: gap of unknown length
 * 35408: contig of 1269 bp in length
 * 35508: gap of unknown length
 * 37018: contig of 1510 bp in length
 * 35509

[illegible]

repeat_region 10708..11008
/note="AluX repeat: matches 1..301 of consensus"
repeat_region 11103..11400
/note="AluX repeat: matches 16..311 of consensus"
repeat_region 11407..11695
/note="AluX repeat: matches 2..290 of consensus"
repeat_region 11809..11951
/note="AluX repeat: matches 161..306 of consensus"
repeat_region 11952..12243
/note="AluX repeat: matches 1..300 of consensus"
repeat_region 12252..12545
/note="AluX repeat: matches 1..307 of consensus"
repeat_region 12553..12854
/note="AluX repeat: matches 1..298 of consensus"
repeat_region 12855..12938
/note="AluX repeat: matches 66..146 of consensus"
repeat_region 13127..13433
/note="AluX repeat: matches 2..308 of consensus"
repeat_region 13882..14004
/note="AluX repeat: matches 14..132 of consensus"
repeat_region 14005..14296
/note="AluX repeat: matches 1..305 of consensus"
repeat_region 14297..14451
/note="AluX repeat: matches 132..300 of consensus"
repeat_region 14486..14790
/note="AluX repeat: matches 1..306 of consensus"
repeat_region 14791..14953
/note="AluX repeat: matches 142..300 of consensus"
repeat_region 14961..15246
/note="AluX repeat: matches 18..299 of consensus"
repeat_region 15707..15972
/note="AluX repeat: matches 1..297 of consensus"
repeat_region 16370..16462
/note="MLTIF repeat: matches 449..541 of consensus"
repeat_region 16654..16945
/note="AluX repeat: matches 1..309 of consensus"
repeat_region 18582..18660
/note="WIR repeat: matches 177..262 of consensus"
repeat_region 18892..19305
/note="L1MC/D repeat: matches 5163..5657 of consensus"
repeat_region 19308..19436
/note="AluX repeat: matches 29..119 of consensus"
repeat_region 19437..19688
/note="AluX repeat: matches 1..252 of consensus"
repeat_region 19762..20061
/note="AluX repeat: matches 1..301 of consensus"
repeat_region 20157..20289
/note="L1M4 repeat: matches 4785..4929 of consensus"
repeat_region 20374..20680
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 21217..21288
/note="WIR repeat: matches 1..72 of consensus"
repeat_region 21574..21708
/note="WIR repeat: matches 63..209 of consensus"
gene complement(join(22679..25420,30411..30572,36490..37690,38418..38471,74872..74936))
/gene="dJ1187M17.1"
complement(join(22679..25420,30411..30572,36490..37690,38418..38471,74872..74936))
/gene="dJ1187M17.1"
/product="dJ1187M17.1.1 (KIAA0860 (isoform1))"
/note="this gene shares its first (5' UTR) exon with gene .5; no overlap of coding regions
match: CDNA8: Em:AB020667
match: ESTs: Em:AW245534 Em:AW919118 Em:BF152150
Em:AW476623 Em:AV605298 Em:BF220115 Em:BE974194
Em:AV588826 Em:BE266831 Em:BE237309 Em:BE485869
Em:AU122539 Em:AU129412 Em:AW525151
/evidence=not experimental
polyA_site complement(22679)
polyA_signal complement(dJ1187M17.1)
complement(22695..22700)

Query Match 18.9%; Score 152; DB 9; Length 125856;
Best Local Similarity 58.7%; Pred. No. 1.2e-15;
Matches 263; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
QY 305 GCAAGCTGGGCACTCGAACAGGACAGACAGGCGCCCTCTGTGTCCTCCCTCCATCT 364
DB 81042 GCCACCTGGGCTCTGGGGAGGCGGAGGTGGAGGCGCTTTCGCGCCCTGCTCACCGC 80983
QY 365 CCACGGACGAGTCAGCATCCAGGAGCTGGAAACAGAACTCTTGGAGAGGAGGCGGCC 424
DB 80982 CCTCCCCAGTGCATCTATCCAGGAGCTGGAGGAGCGGCTTGGAGAGGAGGAGG 80923
QY 425 TCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTTCCTCCAGCTGCGCTACGAGG 484
DB 80922 TGGCAGCTCTGCGCGCAGCCTGAGCAGAGCGAGGCGCTGTGCGCCAGGTACTGGAGG 80863
QY 485 AGCGCGCGCGGCTGCGAGGACGAGCTGGAGGCGCGGAGCGGCGGAGCGGCGGCGGCAACAGC 544
DB 80862 AGCGCGAGAGGCGCTGGAGGCGGAGCTGCGCGAGCTGCGCGAGGCTGCGAGCGGAGC 80803
QY 545 TCAAGCAGGCTCCGACAGAGAGCGCGCGGAGCAGCTGCGCAGTCTGTCACCTGCGAGGTACTGC 604
DB 80802 TACAGCAGTGGCGCGGAGCTGCCAGCGCGCGGAGGAGGCGGCGGCTGATGCGGAGCGG 80743
QY 605 AGCTTCAGCAGGAGAGCGGAGCTCCGCGCAGGAGCTTCAGAGGCTCTATGAAGGAGCAGG 664
DB 80742 GGCTGCGCAGGACAGAAAGCAGCTGCGAGGAGGAGGCGGCGGCTGATGCGGAGCGG 80683
QY 665 ACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGAACAGCTTCGCGCCCGCGC 724
DB 80682 AAGAGCTGGAGGACCAAGGTGCGCGCTGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 80623
QY 725 TGGAGGAGACCCAGTGGGAGGTGAGGCC 752
DB 80622 TAGAGGAACTAAGTGGGAGGTGCGGC 80595
RESULT 13
AB011124
LOCUS AB011124 Homo sapiens mRNA for KIAA0552 protein, complete cds.
DEFINITION AB011124
ACCESSION AB011124.1 GI:3043627
VERSION KIAA0552 protein.
KEYWORDS Homo sapiens male brain cDNA to mRNA, clone lib:pBluescriptII SK
SOURCE plus clone:HH0869.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H.,
Nomura, N. and Ohara, O.
TITLE Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro
DNA Res. 5 (1), 31-39 (1998)
JOURNAL 98290545
MEDLINE 2 (bases 1 to 5257)
REFERENCE Ohara, O., Nagase, T. and Ishikawa, K.
AUTHORS Direct Submission
TITLE Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdnaif@kazusa.or.jp, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES Location/Qualifiers
source 1..5257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HH0869"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
1..5257
gene

Db 183362 AGCAGAGGTGGCAGCTCTGGCGGCGCAGCTGTAGCAGAGCAGCAGCCTGTGGCCCAAG 183303
Qy 476 CCTACGAGGAGCGCGCGCTGCAGGAGAGCTGGAGGGCCGCGAGCCCAAGGCG 535
Db 183302 TGTGGAGAGCGGAGAGGATGGGAGCGGAGCTAGCCAGCTCCGAGAGGCTGCA 183243
Qy 536 GCAACAAGCTCAAGCAGCCTGCGAGAGAGCCAGCGCGCAGCAGCTCTGCACTGCG 595
Db 183242 GTGGGAAGCTGACGAGGAGCGCGCGCTGCGCCAGCGTCCAGCAGGCGCTACAGTGC 183183
Qy 596 AGGTACTGCACTTCAGCAGGAGAGCGCGAGCTCCGCGAGGAGCTCAGAGCCTCATGA 655
Db 183182 AGGTCTCGGTTACAGCAGGACAAAGCTCAGGAGAGCGCGCCAGCTGATAA 183123
Qy 656 AGGACGAGCCTGTGAGAGCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCG 715
Db 183122 GCACAGGAGAGCTGGAGGACAGGTGGCGCTGTTCAGAGGAGGAGCGGCACTTC 183063
Qy 716 GCCCGCGCTGAGAGAGCCAGTGGGAGGTGAGGCC 752
Db 183062 TGCCCCGATGAGGAACTAAGTGGAGGTGCGGCG 183026

RESULT 15

AL772162/c 220355 bp DNA linear HTG 17-AUG-2002
LOCUS Mus musculus chromosome 2 clone RP23-175J8, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.

ACCESSION AL772162
VERSION AL772162.3 GI:22415967
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Almeida, J.

REFERENCE

1 (bases 1 to 220355)
Direct Submission
Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 21, 2002 this sequence version replaced gi:22002755.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm175J8
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 91% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 220027 bases at least Q40
Consensus quality: 220037 bases at least Q30
Consensus quality: 220045 bases at least Q20
Insert size: 220055; sum-of-contigs
Insert size: 215691; 7.3% error; agarose-fp
Quality coverage: 8.70x in Q20 bases; sum-of-contigs Quality
coverage: 8.99x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 152569: contig of 152569 bp in length
* 152570 152669: gap of 100 bp
* 152670 197614: contig of 44945 bp in length

* 197615 197714: gap of 100 bp
* 197715 216175: contig of 18461 bp in length
* 216176 216276: gap of 100 bp
* 216276 220355: contig of 4080 bp in length.
FEATURES
source
1..220355
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-175J8"
/clone_lib="RPCI-23"
1..152569
/note="assembly fragment:02397"
fragment_chain:1
clone_end:T7
vector_side:right
152670..197614
/note="assembly fragment:03756"
fragment_chain:1
197715..216175
/note="assembly fragment:02402"
fragment_chain:1
216276..220355
/note="assembly fragment:03760"
fragment_chain:1
BASE COUNT 58827 a 52680 c 52199 g 56349 t 300 others
ORIGIN
Query_Match 18.4%; Score 148.2; DB 2; Length 220355;
Best Local Similarity 53.7%; Pred. No. 4.4e-15;
Matches 374; Conservative 0; Mismatches 293; Indels 30; Gaps 2;
Qy 86 CTGGCTGTGCTCTGGGGCGCTGTCCAGACTCCGCGCGGAATCCATGTCACAGCTTGCCCA 145
Db 100302 CTGTGGAGTGGGGGGCGCTCTCAGACTCAGCGCGGAATCTACTACAAGCTTGCCCA 100243
Qy 146 CACACAGCACCAGCAGCAGCTACCACTGGACCGCTGGTGCACACCGTGGGAGCCACAA 205
Db 100242 CCTATAGCTCCAGCTATATAGCCAGCATCTGGCACCCCTCAGTGTCTCCACAGCCATATCA 100183
Qy 206 GCGCTTTTGGGGGCTCGCCCAACAT-----CACCCAGG 241
Db 100182 ACCGTATTTGGCACCGCTGGCTATAGTGGCAGCAGCGTGGGGGTGAGGTACCAAG 100123
Qy 242 GCATCGTCTCCAGCAGCAACATGATGACCTGGAAGGCTGTGCTCTCCAGCAGGAG 301
Db 100122 ATCTGGGACCTCTGACAGTGGCGGGCTTCCAGTAAGAGTGGGTGCTCATCATCTATGG 100063
Qy 302 G-----TAGCAAGCTGGGCCACTCGAAACAGGCGAGCAAGGGCCCTCGTGTGTCGCT 355
Db 100062 GGCGGTACAGGCCACTCGGATCCGGGAGGGCGGAAATGGGGCTTGCCATTTGCAGCCT 100003
Qy 356 CCCCCTCTCCAGCAGAGTGCAGCATCCAGGAGCTGGAACAGAGCTGTTGGAGAGG 415
Db 100002 GTTCACCACTCTCCGCCCGCTGATCCAGGAGCTGGAGGAGCGGTGTGGGAGAGG 99943
Qy 416 AGGCGCCCTCCAGAGCTGCAGCGCAGCTTTAGGAGAGAGAGCTTCCCTCCAGCTGG 475
Db 99942 AGCAGAGGTGGAGCTCTGCGGGCGCAGCTTGGAGAGCAGAGCAGCGCTGGCCCCAGG 99883
Qy 476 CCTACGAGGAGCGCGCGCTGTCAGGGACGAGCTGGAGGGCCCGAGAGCCCAAGGCG 535
Db 99882 TGTGGAGGAGCGCGCAGAGGATGGGAGCGGAGCTAGCCGAGCTCCGCGAGGCGCTGCA 99823
Qy 536 GCAACAGCTCAAGCAGGCTCGCAGAGAGAGCCAGCGCGCGCAGAGGTCTCTGCACCTGC 595
Db 99822 GTGGGAAGCTGCAGCAGGTGGCGCGCTGTCAGCGCTGCCAGCAGGCGCTACAGCTGC 99763
Qy 596 AGGTACTGCACTTCAGCAGGAGAGCGGAGCTCCCGCAGAGCTCCAGAGCCTCATGA 655
Db 99762 AGGTGCTCGGTTACAGCAGGAGCAAGAAACAGCTGCAGGAGGAGCGCGCCCGCTGATAA 99703
Qy 656 AGGAGCAGGACCTCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGAGACCAAGCTTCG 715

Db 99702 GGCAACGGGAAGAGCTGGAGGACAAAGGTGGCCGTCTGTCAAGAGGAGCGCGACTTCC 99643
Qy 716 GCCCCGCGCTGGAGGAGACCCAGTGGGAGGTGAGGCC 752
Db 99642 TCCCCCGGATGGAGGAAACTAAGTGGGAGGTCCGGGC 99606

Search completed: June 15, 2003, 02:18:36
Job time : 2097.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:38:41 ; Search time 185.106 Seconds
(without alignments)
9781.434 Million cell updates/sec

Title: US-09-513-888C-1_COPY_1707_2510
Perfect score: 804
Sequence: 1 gccatcctgactctcccc.....tttggccagatccccctt 804

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				
N_Geneseq 101002:*				
1:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*			
2:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*			
3:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*			
4:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*			
5:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*			
6:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*			
7:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*			
8:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*			
9:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*			
10:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*			
11:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*			
12:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*			
13:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*			
14:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*			
15:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*			
16:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*			
17:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*			
18:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*			
19:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*			
20:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*			
21:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*			
22:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*			
23:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*			
24:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	804	100.0	9048	21	AAA64507
2	802.4	99.8	2411	22	AAA30637
3	802.4	99.8	2411	22	AAA28699
4	746.2	92.8	1512	21	AAA64513
5	746.2	92.8	1614	21	AAA64512
6	746.2	92.8	1692	21	AAA64514
7	746.2	92.8	1722	21	AAA64515
8	746.2	92.8	1791	21	AAA64509
9	746.2	92.8	5492	21	AAA64508

10	463.4	57.6	560	22	AAA30595	DNA encoding novel
11	463.4	57.6	560	22	AAA28127	Novel cDNA encoding
12	157.6	19.6	633	21	AAA64511	Nucleotide sequenc
c 13	112.4	14.0	561	24	ABK62287	Rat sequence diffe
14	82.8	10.3	795	19	AAV55830	FLGA insert stabl
15	82.8	10.3	799	19	AAV55831	Nucleotide sequenc
16	82.8	10.3	1926	21	AAA50254	Epstein Barr virus
17	82.8	10.3	1926	22	AAAF82902	EBV tethering prot
18	82.8	10.3	2580	21	AAA75454	Nucleotide sequenc
19	82.8	10.3	2580	24	AAI64275	Epstein-Barr virus
c 20	82.8	10.3	5452	20	AAAX90923	Anti-sense strand
c 21	82.8	10.3	8705	20	AAZ23778	Vector shuttle DN
22	82.8	10.3	9600	19	AAV21683	Vector plasmid pCM
23	82.8	10.3	10380	20	AAZ22248	Nucleotide sequenc
24	82.8	10.3	10596	14	AAQ51731	Plasmid pCisEBON f
25	82.8	10.3	10596	17	AAT40348	Plasmid pCisEBON f
26	82.8	10.3	10596	20	AAI15650	Nucleotide sequenc
c 27	82.8	10.3	16080	21	AAA59553	DNA clone pCEK Cl.
28	79.6	9.9	1623	24	ABQ54361	Human ovarian anti
29	77.2	9.6	1925	20	AAAX90924	Epstein Barr Virus
30	77.2	9.6	5120	22	AAC84677	DNA sequence of hu
c 31	68.8	8.6	16442	18	AAAX83006	Partial mouse WRN
32	67	8.3	1617	21	AAFI5625	Human prostate can
33	67	8.3	2108	24	ABL67774	Oesophagus cancer
34	67	8.3	3489	21	AAA30290	Kaposi's sarcoma-a
35	67	8.3	3489	22	AAAF82901	Nucleotide sequenc
36	67	8.3	3489	24	ABA93487	Kaposi's sarcoma-a
c 37	67	8.3	32207	20	AAV73805	KSHV LUR DNA (nucl
c 38	67	8.3	137507	19	AAV19941	KSHV long unique c
39	64.6	8.0	9551	20	AAZ22301	CDNA encoding a hu
c 40	62.2	7.7	1159	21	AAA59240	An EcoRI fragment
41	61.8	7.7	477	21	AAZ44307	Human SCA7 genomic
c 42	61.6	7.7	1037	21	AAA59242	Exons E, C and A o
c 43	61.6	7.7	1472	21	AAA59241	Exons D, C, B and
44	60	7.5	1416	23	AAAS71051	DNA encoding novel
45	60	7.5	4720	24	AAAS94782	Human DNA sequence

ALIGNMENTS

RESULT 1

AAA64507

ID AAA64507 standard; DNA; 9048 BP.

AC AAA64507;

XX

XX 02-JAN-2001 (first entry)

DT

XX Nucleotide sequence comprising the human FEZ1 gene.

DE

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;

KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;

KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;

KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;

KW tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.

XX

XX WO2000050565-A2.

PN

XX

XX 31-AUG-2000.

PD

XX

XX 25-FEB-2000; 2000WO-US04950.

PF

XX

XX 25-FEB-1999; 99US-0121537.

PR

XX (UYJE-) UNIV JEFFERSON THOMAS.

PA

XX Croce CM, Ishii H;

PI

XX WPI; 2000-558396/51.

DR

XX


```
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.

PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476224/51.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the respiratory system including respiratory
XX cancers and also for testing and detection e.g. diagnosis -
XX
XX Disclosure; SED ID No 1133; 546pp; English.
XX
XX The present invention relates to the isolation of novel human
XX respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
XX sequences encoding for these polypeptides. The sequences of the
XX invention are useful for preventing, treating and/or prognosing
XX disorders related to the respiratory system including throat
XX disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
XX lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
XX pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
XX the respiratory tissues e.g. lung cancer. The polynucleotide sequences
XX of the invention are useful in gene therapy and antisense therapy.
XX
XX AAS28161-AAS28764 represent genomic sequences encoding for novel
XX human respiratory antigens.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2411 BP; 558 A; 681 C; 733 G; 439 T; 0 other;

Query Match 99.8%; Score 802.4; DB 22; Length 2411;
Best Local Similarity 99.9%; Pred. No. 1.9e-142;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCAGCTGACACCGCCGCGCTTCCAGACTCCGAC 60
Db |||||
Qy 1367 GCCATCTGCACTCTCCCGGAGAGTGCCAGCAGCTGACACCGCCGCGCTTCCAGACTCCGAC 1426
Db |||||
Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTGCTTGGGGCGCTGTTCAGACTCCGCGC 120
Db |||||
Qy 1427 AAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTGCTTGGGGCGCTGTTCAGACTCCGCGC 1486
Db |||||
Qy 121 CGGAATCTCCATGTCCAGCTGCGCCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180
Db |||||
Qy 1487 CGGAATCTCCATGTCCAGCTGCGCCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1546
Db |||||
Qy 181 CTGGTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCCGCCCAACATCACCAGCAG 240
Db |||||
Qy 1547 CTGGTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCCGCCCAACATCACCAGCAG 1606
Db |||||
Qy 241 GGCAATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACCGGA 300
Db |||||
Qy 1607 GGCAATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACCGGA 1666
Db |||||
Qy 301 GGTAGCAAGCTGGGCGCACTCGAAACAAGCAGACAAAGGGCCCTCGTGTGTCGCTCCGCC 360
Db |||||
Qy 1667 GGTAGCAAGCTGGGCGCACTCGAAACAAGCAGACAAAGGGCCCTCGTGTGTCGCTCCGCC 1726
Db |||||
Qy 361 ATCTCCAGGACGAGTGCAGCATCCAGGATCCAGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 420
Db |||||
Qy 1727 ATCTCCAGGACGAGTGCAGCATCCAGGATCCAGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 1786
Db |||||
Qy 421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTTCAGGCTTGCCCTTAC 480
Db |||||
```

Db 1787 GCCCTCCAGAAAGCTGCAGCGCAGCTTTGAGGAGAAAGGAGCTTGCTCCAGCCTGCGCCTAC 1846
Qy 481 GAGGAGCGCGCGCGCTGCAGGACAGAGCTGAGAGCGCGGAGCCCAAGCGCGCAAC 540
Db 1847 GAGGAGCGCGCGCGCTGCAGGACAGAGCTGAGAGCGCGGAGCCCAAGCGCGCAAC 1906
Qy 541 AAGCTCAAGCAGGCTTCGAGAAAGCAGCGCGCGCAGCGCTGCACTGCAAGTA 600
Db 1907 AAGCTCAAGCAGGCTTCGAGAAAGCAGCGCGCGCAGCGCTGCACTGCAAGTA 1966
Qy 601 CTGCGAGCTTCAGCAGGAGCGCAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 660
Db 1967 CTGCGAGCTTCAGCAGGAGCGCAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 2026
Qy 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAAACAGAGCTTCGCGCCC 720
Db 2027 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAAACAGAGCTTCGCGCCC 2086
Qy 721 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTTGGGTGCTCA 780
Db 2087 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTTGGGTGCTCA 2146
Qy 781 CGGCTTGGCGCCAGTACCCCT 804
Db 2147 CGGCTTGGCGCCAGTACCCCT 2170

RESULT 4

AAA64513

ID AAA64513 standard; cDNA; 1512 BP.

XX AC AAA64513;

XX DT

XX 02-JAN-2001 (first entry)

XX DE

XX Nucleotide sequence of truncated FEZ1 transcript D14.

XX Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1512

XX /tag= a

XX /product= "truncated FEZ1"

XX WO200050565-A2.

XX PN

XX 31-AUG-2000.

XX PD

XX 25-FEB-2000; 2000WO-US04950.

XX PF

XX 25-FEB-1999; 99US-0121537.

XX PR

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX PA

XX Croce CM, Iehli H;

XX PI

XX WPI; 2000-558396/51.

XX DR

XX P-PSDB; AAB08720.

XX XX

XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -

XX PS

XX Disclosure; Fig 5F; 255pp; English.

XX XX

XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour

CC suppressor gene, located at chromosome location 9p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.

XX SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;

Query Match

Best Local Similarity 92.8%; Score 746.2; DB 21; Length 1512;

Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACCAAGCTGCAACCCGCTCCAGAC 60

Db 406 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACCAAGCTGCAACCCGCTCCAGAC 465

Qy 61 AAGCCCAAGGAGGAGGAGTGAAGCTGGCTGTCTTGGGGCGCTTCAGACTCCGCG 120

Db 466 AAGCCCAAGGAGGAGGAGTGAAGCTGGCTGTCTTGGGGCGCTTCAGACTCCGCG 525

Qy 121 CGGAATCTCATGTCCAGCTTCCCAACACACAGCAGCAGCAGCTACAGCTGACCCG 180

Db 526 CGGAATCTCATGTCCAGCTTCCCAACACACAGCAGCAGCAGCTACAGCTGACCCG 585

Qy 181 CTGTCTCACACCCGTGGGACCCCAAGCGCTTTGGGGCTCCGCCCAACATCAACCCAG 240

Db 586 CTGTCTCACACCCGTGGGACCCCAAGCGCTTTGGGGCTCCGCCCAACATCAACCCAG 645

Qy 241 GGCATCTCTCCAGGACAGCAATGATGAGCTCTGAAGGCTCTGTCTTCTCGACCGA 300

Db 646 GGCATCTCTCTCCAGGACAGCAATGATGAGCTCTGAAGGCTCTGTCTTCTCGACCGA 705

Qy 301 GGTAGCAAGCTGGGCGCACTCGAACAAGGCACACAAGGSCCTCTGTGTCTCGCTCCCCC 360

Db 706 GGTAGCAAGCTGGGCGCACTCGAACAAGGCACACAAGGSCCTCTGTGTCTCGCTCCCCC 765

Qy 361 ATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGCG 420

Db 766 ATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGCG 825

Qy 421 GCCCTCAGAAAGCTGCAGCGCAGCTTTGAGGAGAAAGAGCTTGCTCCAGCTTGCCTTAC 480

Db 826 GCCCTCAGAAAGCTGCAGCGCAGCTTTGAGGAGAAAGAGCTTGCTCCAGCTTGCCTTAC 885

Qy 481 GAGGAGCGCGCGCGCTGCAGGAGCAGAGCTGGAGGSCCGGAGCCCAAGGCGGCAAC 540

Db 886 GAGGAGCGCGCGCGCTGCAGGAGCAGAGCTGGAGGSCCGGAGCCCAAGGCGGCAAC 945

Qy 541 AAGCTCAAGCAGGCTTCGAGAAAGCAGCGCGCAGCAGAGGTCCTGACCTGCAGGTA 600

Db 946 AAGCTCAAGCAGGCTTCGAGAAAGCAGCGCGCAGCAGAGGTCCTGACCTGCAGGTA 1005

Qy 601 CTGAGCTTCAGCAGGAGAGCGGCGAGCTCCGCGAGAGCTTCGAGAGCTCATGAGGAG 660

Db 1006 CTGAGCTTCAGCAGGAGAGCGGCGAGCTCCGCGAGAGCTTCGAGAGCTCATGAGGAG 1065

Qy 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTTACAGAGGAGGAGAACAGCTTCGCGCCC 720

Db 1066 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTTACAGAGGAGGAGAACAGCTTCGCGCCC 1125

Qy 721 CGCTGGAGGAGACCCAGTGGGAGGTGAGGC 751

FT	CDS	1..1692	
FT		/*tag= a	
FT		/product= "truncated FEZ1"	
XX			
PN		WO200050565-A2.	
XX			
PD		31-AUG-2000.	
XX			
PF		25-FEB-2000; 2000WO-US04950.	
XX			
XX		25-FEB-1999; 99US-0121537.	
XX			
PA		(UYJE-) UNIV JEFFERSON THOMAS.	
XX			
PI		Croce CM, Ishii H;	
XX			
XX		WPI; 2000-558396/51.	
DR		P-PSDB; AAB08721.	
XX			
PT	New polynucleotide homologous with a portion of one strand of the human		
PT	FEZ1 gene, useful for alleviating abnormal cell proliferation such as		
PT	cancer -		
XX			
PS	Disclosure; Fig 5G; 255pp; English.		
XX			
CC	The present sequence encodes a truncated human FEZ1 polypeptide. The		
CC	encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour		
CC	suppressor gene, located at chromosome location 8p22. Decreased		
CC	or no expression of FEZ1 is detected in a variety of cancer cells.		
CC	Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1		
CC	also interacts with tubulin, with microtubules, and with protein		
CC	EFl-gamma. Post-translational phosphorylation and dephosphorylation		
CC	modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene		
CC	expression are useful for inducing cells to proliferate. Compounds		
CC	which modulate FEZ1 association with tubulin are useful for alleviating		
CC	tubulin hyper- or hypo- polymerisation disorders, such as those		
CC	associated with aberrant initiation of mitosis, modulation of the		
CC	initiation and rate of cell proliferation and cell growth, modulation of		
CC	cell shape, cell rigidity, cell motility, rate and stage of cellular		
CC	DNA replication, intracellular distribution of organelles, metastatic		
CC	potential of cell and cellular transformation from a non-cancerous to		
CC	cancerous phenotype. Compounds which modulate FEZ1 binding and		
CC	phosphorylation are also useful for alleviating a disorder, such as		
CC	tumorigenesis, tumour survival, growth and metastasis.		
XX			
SQ	Sequence 1692 BP; 380 A; 523 C; 236 T; 0 other;		
Query Match 92.8%; Score 746.2; DB 21; Length 1692;			
Best Local Similarity 99.6%; Pred. No. 7.1e-132;			
Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	1	GCATCTGCACTCTCCCGGAGAGTGCCAGCCACGAGTGCACCCGCCCTCCAGAC	60
DB	406	GCATCTGCACTCTCTCCCGGAGAGTGCCAGCCACGAGTGCACCCGCCCTCCAGAC	465
QY	61	AAGCCCAAGGAGCAGGAGCTGAAGCTGGCGCTGTGCTGGGGCGCTGTGAGACTCCGGC	120
DB	466	AAGCCCAAGGAGCAGGAGCTGAAGCTGGCGCTGTGCTGTGGGGCGCTGTGAGACTCCGGC	525
QY	121	CGGAATCATGTCCAGCTGCCACACAGCACAGCAGCAGCAGCTACCACTGGACCCG	180
DB	526	CGGAATCATGTCCAGCTGCCACACAGCACAGCAGCAGCAGCTACCACTGGACCCG	585
QY	181	CTGGTTCACCCGCTGGGACCCACAGCCGTTTTGGGGGCTCCGCCACACATACCCAG	240
DB	586	CTGGTTCACCCGCTGGGACCCACAGCCGTTTTGGGGGCTCCGCCACACATACCCAG	645
QY	241	GCATCTGCTCTCCAGCAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGGA	300
DB	646	GCATCTGCTCTCCAGCAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGGA	705
QY	301	GGTAGCAAGCTGGGGCACTTCGAACAGGCAGCAAGGGGCCCTCTGTGTCCGCTCCCCC	360

Db	706	GGTAGCAAGCTGGGGCACTTCGAACAGGCAGACAGGGCCCCCTCTGTGTCCGCTCCCCC	765
QY	361	ATCTCCACGAGCAGAGTGAGCATCCAGCAGCTTGAACAGAAAGCTTTGGAGAGGAGGCG	420
Db	766	ATCTCCACGAGCAGAGTGAGCATCCAGCAGCTTGAACAGAAAGCTTTGGAGAGGAGGCG	825
QY	421	GCCTTCCAGAAAGCTGCAGCGCAGCTTTGAGGAGAAAGAGCTTGCCTCCAGCCTTGGCCTAC	480
Db	826	GCCTTCCAGAAAGCTGCAGCGCAGCTTTGAGGAGAAAGAGCTTGCCTCCAGCCTTGGCCTAC	885
QY	481	GAGGAGCGCGCGCGCTGCAGGAGCAGCTGGAGGCGCGGAGCCCAAGGCGGCGCAAC	540
Db	886	GAGGAGCGCGCGCGCTGCAGGAGCAGCTGGAGGCGCGGAGCCCAAGGCGGCGCAAC	945
QY	541	AAGCTCAAGCAGGCGCTTCGAGAAGAGCCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA	600
Db	946	AAGCTCAAGCAGGCGCTTCGAGAAGAGCCAGCGCGCAGCAGGCTCTGCACCTGCAGGTA	1005
QY	601	CTGAGCTTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAG	660
Db	1006	CTGAGCTTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAG	1065
QY	661	CAGGACCTGCTGGAGAGCAAGCTCAGGTCCTACGAGAGGAGAGACAGCTTCCGGCCCC	720
Db	1066	CAGGACCTGCTGGAGAGCAAGCTCAGGTCCTACGAGAGGAGAGACAGCTTCCGGCCCC	1125
QY	721	GCCTGGAGGAGAGCCAGTGGGAGGTGAGGC	751
Db	1126	GCCTGGAGGAGAGCCAGTGGGAGGTGAGGC	1156

RESULT 7			
AAA64515			
ID	AAA64515 standard; cDNA; 1722 BP.		
XX	AC AAA64515;		
XX	02-JAN-2001 (first entry)		
DE	Nucleotide sequence of truncated FEZ1 transcript G3612.		
KW	Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;		
KW	tumour proliferation; tubulin; microtubule; protein EFl-gamma;		
KW	tubulin polymerisation disorder; mitosis initiation; cell proliferation;		
KW	cell growth; cell shape; cell rigidity; cell motility; DNA replication;		
KW	tumorigenesis; tumour survival; metastasis; ss.		
OS	Homo sapiens.		
XX	Key Location/Qualifiers		
FT	1..1722		
FT	/*tag= a		
FT	/product= "truncated FEZ1"		
XX	WO200050565-A2.		
XX	31-AUG-2000.		
XX	25-FEB-2000; 2000WO-US04950.		
XX	25-FEB-1999; 99US-0121537.		
XX	(UYJE-) UNIV JEFFERSON THOMAS.		
XX	Croce CM, Ishii H;		
XX	WPI; 2000-558396/51.		
XX	P-PSDB; AAB08722.		
XX	New polynucleotide homologous with a portion of one strand of the human		
XX	FEZ1 gene, useful for alleviating abnormal cell proliferation such as		
XX	cancer -		

Db 877 ATCTCCAGCAGCTGCGAGCATCCAGGAGCTGGAGCAGAAAGCTGTTGGAGAGGGAGGC 936
Qy 421 GCCCTCCAGAGCTCCAGCGAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGCGCTAC 480
Db 937 GCCCTCCAGAGCTGCGAGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGCGCTAC 996
Qy 481 GAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGGAGCCCAAGGGGGCAAC 540
Db 997 GAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGGAGCCCAAGGGGGCAAC 1056
Qy 541 AAGCTCAAGAGGCTCCAGAGAGCAGCGCGCGAGAGGCTTGCTCCAGCTGCGAGTA 600
Db 1057 AAGCTCAAGAGGCTCCAGAGAGCAGCGCGCGAGAGGCTTGCTCCAGCTGCGAGTA 1116
Qy 601 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 660
Db 1117 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 1176
Qy 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGAGGGAGAGACAGCTTCGGCCCC 720
Db 1177 CAGGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGAGGGAGAGACAGCTTCGGCCCC 1236
Qy 721 GCGCTGGAGGAGACCCAGTGGAGGTGAGGC 751
Db 1237 GCGCTGGAGGAGACCCAGTGGAGGTGAGGC 1267

RESULT 10

AA30595
ID AA30595 standard; cDNA; 560 BP.

XX AA30595;

XX 21-NOV-2001 (first entry)

DE DNA encoding novel lung cancer antigen, Seq ID No 17.

XX Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasm;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; skin aging;
KW ocular disorder; wound healing; organ transplantation; ss.

XX Homo sapiens.

XX WO200155300-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-0501238.

XX 31-JAN-2000; 2000US-0179065.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465565/50.

XX P-PSDB; AAU18935.

PT Isolated nucleic acid molecule encoding a lung cancer antigen is used
PT in preventing, treating or ameliorating a medical condition -

XX Claim 1; SEQ ID No 17; 475pp; English.

XX The invention relates to novel isolated lung cancer antigen
CC polynucleotides (I) and polypeptides (II). (I) and (II) are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are
CC also used in diagnosing a pathological condition or susceptibility to a
CC pathological condition, in particular, lung cancer. The antibodies to

CC (II) can also be used in alleviating symptoms associated with the
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC treated include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. The polypeptides can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. AA30595-AAS30685 represent novel human lung cancer antigen
CC coding sequences, PCR primers and related sequences of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 560 BP; 120 A; 196 C; 161 G; 83 T; 0 other;

Query Match 57.68; Score 463.4; DB 22; Length 560;
Best Local Similarity 99.8%; Pred. No. 1.3e-78;

Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTCGACCTCTCCCGGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC 60

Db 96 GCCATCTCGACCTCTCCCGGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC 155

Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTCTGGGGCGCTGTGAGATCCGGC 120

Db 156 AAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTCTGGGGCGCTGTGAGATCCGGC 215

Qy 121 CGGAATCTCCATGTCCAGCCTGCCACACAGCAGCAGCAGCTACCGAGCCCG 180

Db 216 CGGAATCTCCATGTCCAGCCTGCCACACAGCAGCAGCAGCTACCGAGCCCG 275

Qy 181 CTGTGTACACCCGTTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACATCACCAG 240

Db 276 CTGTGTACACCCGTTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACATCACCAG 335

Qy 241 GGATCTCTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGAGGA 300

Db 336 GGATCTCTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGAGGA 395

Qy 301 GGTAGCAAGCTGGGCGCTCGAACAAGCAGCAGCAGCAGCAGCTGCTGCTCCCGC 360

Db 396 GGTAGCAAGCTGGGCGCTCGAACAAGCAGCAGCAGCAGCTGCTGCTCCCGC 455

Qy 361 ATCTCCAGCAGCAGTGCAGCATCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGC 420

Db 456 ATCTCCAGCAGCAGTGCAGCATCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGC 515

Qy 421 GCCCTCCAGAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCC 465

Db 516 GCCCTCCAGAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCC 560

RESULT 11

AA328127
ID AA328127 standard; cDNA; 560 BP.

XX AA328127;

XX 07-NOV-2001 (first entry).

XX Novel cDNA encoding for human respiratory antigen #259.

XX Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti-allergic; anti-asthmatic; anti-inflammatory; olfactory;
KW respiratory active; ss.
XX Homo sapiens.

XX WO200155448-A1.
PN 27-SEP-2000; 2000US-0235836.
XX 29-SEP-2000; 2000US-0236327.
PD 29-SEP-2000; 2000US-0236367.
XX 29-SEP-2000; 2000US-0236368.
XX 29-SEP-2000; 2000US-0236369.
PF 29-SEP-2000; 2000US-0236370.
XX 29-SEP-2000; 2000US-0236802.
XX 02-OCT-2000; 2000US-0237037.
XX 02-OCT-2000; 2000US-0237038.
XX 02-OCT-2000; 2000US-0237039.
XX 02-OCT-2000; 2000US-0237040.
XX 13-OCT-2000; 2000US-0239935.
XX 13-OCT-2000; 2000US-0239937.
XX 20-OCT-2000; 2000US-0240960.
XX 20-OCT-2000; 2000US-0241221.
XX 20-OCT-2000; 2000US-0241785.
XX 20-OCT-2000; 2000US-0241786.
XX 20-OCT-2000; 2000US-0241787.
XX 20-OCT-2000; 2000US-0241808.
XX 20-OCT-2000; 2000US-0241809.
XX 20-OCT-2000; 2000US-0241826.
XX 01-NOV-2000; 2000US-0244617.
XX 08-NOV-2000; 2000US-0246474.
XX 08-NOV-2000; 2000US-0246475.
XX 08-NOV-2000; 2000US-0246476.
XX 08-NOV-2000; 2000US-0246477.
XX 08-NOV-2000; 2000US-0246478.
XX 08-NOV-2000; 2000US-0246523.
XX 08-NOV-2000; 2000US-0246524.
XX 08-NOV-2000; 2000US-0246525.
XX 08-NOV-2000; 2000US-0246526.
XX 08-NOV-2000; 2000US-0246527.
XX 08-NOV-2000; 2000US-0246528.
XX 08-NOV-2000; 2000US-0246532.
XX 08-NOV-2000; 2000US-0246609.
XX 08-NOV-2000; 2000US-0246610.
XX 08-NOV-2000; 2000US-0246611.
XX 08-NOV-2000; 2000US-0246613.
XX 17-NOV-2000; 2000US-0249207.
XX 17-NOV-2000; 2000US-0249208.
XX 17-NOV-2000; 2000US-0249209.
XX 17-NOV-2000; 2000US-0249210.
XX 17-NOV-2000; 2000US-0249211.
XX 17-NOV-2000; 2000US-0249212.
XX 17-NOV-2000; 2000US-0249213.
XX 17-NOV-2000; 2000US-0249214.
XX 17-NOV-2000; 2000US-0249215.
XX 17-NOV-2000; 2000US-0249216.
XX 17-NOV-2000; 2000US-0249217.
XX 17-NOV-2000; 2000US-0249218.
XX 17-NOV-2000; 2000US-0249244.
XX 17-NOV-2000; 2000US-0249245.
XX 17-NOV-2000; 2000US-0249284.
XX 17-NOV-2000; 2000US-0249285.
XX 17-NOV-2000; 2000US-0249297.
XX 17-NOV-2000; 2000US-0249299.
XX 17-NOV-2000; 2000US-0249300.
XX 01-DEC-2000; 2000US-0250160.
XX 01-DEC-2000; 2000US-0250391.
XX 05-DEC-2000; 2000US-0251030.
XX 05-DEC-2000; 2000US-0251988.
XX 05-DEC-2000; 2000US-0256719.
XX 06-DEC-2000; 2000US-0251479.
XX 08-DEC-2000; 2000US-0251866.
XX 08-DEC-2000; 2000US-0251869.
XX 08-DEC-2000; 2000US-0251989.
XX 08-DEC-2000; 2000US-0251990.
XX 11-DEC-2000; 2000US-0254097.
XX 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476224/51.
DR P-PSDB; AAU17943.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the respiratory system including respiratory
PT cancers and also for testing and detection e.g. diagnosis -
XX
XX Claim 4; SED ID No 269; 546pp; English.
PS
XX The present invention relates to the isolation of novel human
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful for preventing, treating and/or prognosing
CC disorders related to the respiratory system including throat
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS27869-AAS28159 encode for novel human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 560 BP; 120 A; 196 C; 161 G; 83 T; 0 other;
PS
Query Match 57.6%; Score 463.4; DB 22; Length 560;
Best Local Similarity 99.8%; Pred. No. 1.3e-78;
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 96 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 155
QY 61 AAGCCCAAGGAGCAGAGCTGAAGCTGGCGTGTCTGTGGGGCTGTCTGAGACTCCGGC 120
Db 156 AAGCCCAAGGAGCAGAGCTGAAGCTGGCGTGTCTGTGGGGCTGTCTGAGACTCCGGC 215
QY 121 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACCGCTGGACCCG 180
Db 216 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACCGCTGGACCCG 275
QY 181 CTGGTACACCGCTGGGACCCACAGCCGTTTTGGGGGCTCGGCCACAAATACCCAG 240
Db 276 CTGGTACACCGCTGGGACCCACAGCCGTTTTGGGGGCTCGGCCACAAATACCCAG 335
QY 241 GCATCGTCTCCAGGACGACCAATGATGAGCCTGAAGGCTCTGCTTCTCCGACGGA 300
Db 336 GCATCGTCTCCAGGACGACCAATGATGAGCCTGAAGGCTCTGCTTCTCCGACGGA 395
QY 301 GGTAGCAAGCTGGGCACTTCGAACAAGGACGACAAGGGCCCTCTGTGTGCGTCCCCC 360
Db 396 GGTAGCAAGCTGGGCACTTCGAACAAGGACGACAAGGGCCCTCTGTGTGCGTCCCCC 455
QY 361 ATCTCCAGGACGAGTGCAGATCCAGAGCTGGACAGAGCTTTGGAGAGGAGGCG 420
Db 456 ATCTCCAGGACGAGTGCAGATCCAGAGCTGGACAGAGCTTTGGAGAGGAGGCG 515
QY 421 GCCCTCCAGAGCTGCAGGCGAGCTTTGAGGAGAGGAGGCTTGCC 465
Db 516 GCCCTCCAGAGCTGCAGGCGAGCTTTGAGGAGAGGAGGCTTGCC 560
RESULT 12
ID AAA64511
XX AAA64511 standard; cDNA; 633 BP.
XX
AC AAA64511;
XX
XX 02-JAN-2001 (first entry)
XX

DE Nucleotide sequence of truncated FEZ1 transcript E264162.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein E1-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..633
FT /tag= a
FT /product= "truncated FEZ1"
XX
XX MO200050565-A2.
XX 31-AUG-2000.
XX 25-FEB-2000; 2000WO-US04950.
XX 25-FEB-1999; 99US-0121537.
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX Croce CM, Ishii H;
XX WPI; 2000-558396/51.
XX P-PSDB; AAB08718.
XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
XX Disclosure; Fig 5D; 255pp; English.
XX
XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC E1-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 633 BP; 136 A; 217 C; 175 G; 105 T; 0 other;
Query Match 19.6%; Score 157.6; DB 21; Length 633;
Best Local Similarity 92.2%; Pred. No. 6e-21;
Matches 166; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 406 GCCATCTGCACTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 465
QY 61 AAGCCCAAGGAGCAGAGCTGAAGCTGGCGTGTCTGTGGGGCTGTCTGAGACTCCGGC 120
Db 466 AAGCCCAAGGAGCAGAGCTGAAGCTGGCGTGTCTGTGGGGCTGTCTGAGACTCCGGC 525
QY 121 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACCGCTGGACCCG 180
Db 526 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTTCGAGGTTGACCTG 585

RESULT 13
ABK62287/c
ID ABK62287 standard; cDNA; 561 BP.
XX
AC ABK62287;
XX
AT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #194.
XX
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
PN WO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US23872.
XX
PR 31-JUL-2000; 2000US-222040P.
PR 02-NOV-2000; 2000US-244880P.
PR 11-MAY-2001; 2001US-290029P.
PR 15-MAY-2001; 2001US-290645P.
PR 22-MAY-2001; 2001US-292336P.
PR 06-JUN-2001; 2001US-295798P.
PR 13-JUN-2001; 2001US-297457P.
PR 19-JUN-2001; 2001US-298884P.
PR 09-JUL-2001; 2001US-303459P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
DR WPI; 2002-241625/29.
XX
PT Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or
PT cells exposed to the toxin and comparing these to gene expression in
PT unexposed tissues or cells -
XX
PS Claim 1; Seq ID No 194; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic
CC effect of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression.
CC The method can also be used to identify an agent which modulates the
CC toxic response and predict cellular pathways that a compound modulates
CC in a cell. The methods utilise a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity

CC is characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent.
XX
SQ Sequence 561 BP; 151 A; 143 C; 127 G; 140 T; 0 other;
Query Match 14.0%; Score 112.4; DB 24; Length 561;
Best Local Similarity 77.0%; Pred. No. 2e-12;
Matches 137; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 367 ACGGACGAGTCAGCATCCAGGAGCTGGAAACAGAGCTGTTCGAGAGGAGGGCCCTC 426
DB 561 ACGGATGAGTCACCAATTCAGGAGCTGGAGCAGAGCTGTTCGAGCGGAGACTGCACTA 502
QY 427 CAGAAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTGCCTCCAGCTGCCTACGAGGAG 486
DB 501 CAGAAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTGCCTCCAGCTGCCTACGAGGAG 442
QY 487 CGGCGGCGGCTGCAGGAGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 544
DB 441 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 384
RESULT 14
AAV55830
ID AAV55830 standard; DNA; 795 BP.
XX
AC AAV55830;
XX
AT 18-NOV-1998 (first entry)
XX
DE FLGA insert stabilising polypeptide encoding DNA.
XX
KW Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.
XX
OS Epstein-barr virus.
XX
FH Key Location/Qualifiers
FT CDS 1..788
FT FT /*tag= a
FT FT /product= "stabilising polypeptide"
XX
PN WO9822577-A1.
XX
PD 28-MAY-1998.
XX
PF 17-NOV-1997; 97WO-IB01508.
XX
PR 25-JUN-1997; 97US-0048945.
PR 15-NOV-1996; 96US-0030986.
XX
PA (MASU/) MASUCCI M G.
XX
PI Masucci MG;
XX
DR WPI; 1998-312463/27.
XX P-PSDB; AAW79128.
XX
PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
XX a peptide sequence containing glycine repeats
PS Disclosure; Fig 3; 120pp; English.
XX
CC This DNA encodes a stabilising polypeptide and is the FLGA insert of the
CC invention. The invention provides a method for increasing the resistance
CC of a core protein to proteolytic degradation that comprises linking or
CC inserting onto or into the core protein a stabilising polypeptide of
CC formula [(Glya)x(Glyb)y(Glyc)z]n where Glya, Glyb, Glyc are 1-6

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:27:06 ; Search time 47.1752 Seconds
(without alignments)
5226.648 Million cell updates/sec

Title: US-09-513-888C-1_COPY_1707_2510

Perfect score: 804

Sequence: 1 gccatcctgactctcccc.....tttggccagtagtaccctccct 804

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.8	10.3	1926	4	US-09-249-585A-2
2	82.8	10.3	2580	3	US-09-050-863-2
3	82.8	10.3	2580	4	US-09-359-081-2
4	82.8	10.3	5452	2	US-09-130-114-1
5	82.8	10.3	9600	4	US-08-910-647-1
6	82.8	10.3	9600	4	US-09-620-925-1
7	82.8	10.3	10596	1	US-07-884-811-15
8	82.8	10.3	10596	1	US-07-885-971-15
9	82.8	10.3	10596	1	US-08-087-783A-15
10	82.8	10.3	10596	1	US-08-194-088B-15
11	82.8	10.3	10596	2	US-08-194-087-15
12	82.8	10.3	10596	5	PCT-US93-04648-15
13	68.8	8.6	16442	3	US-08-781-891-208
14	67	8.3	3489	2	US-08-728-323A-1
15	67	8.3	3489	4	US-09-298-568-1
16	67	8.3	32207	2	US-08-770-379-20
17	67	8.3	32207	4	US-08-757-669A-20
18	67	8.3	32207	4	US-09-230-371A-20
19	64.6	8.0	9551	1	US-08-056-200-93
20	64.6	8.0	9551	2	US-08-800-644-93
21	61.8	7.7	477	4	US-09-135-994-1
22	59.8	7.4	397	3	US-09-253-691-3
23	59.4	7.4	234	2	US-08-469-802B-3
24	59.4	7.4	234	2	US-08-267-803B-3
25	56.4	7.0	2793	1	US-08-209-747-1
26	56.4	7.0	2793	1	US-08-458-298-1
27	56	7.0	195	1	US-08-469-802B-2

28	56	7.0	195	2	US-08-267-803B-2	Sequence 2, Appli
29	55.4	6.9	543	6	5273901-6	Patent No. 5273901
30	54.4	6.8	2313	4	US-09-370-838-157	Sequence 157, App
31	53.8	6.7	168	1	US-08-469-802B-4	Sequence 4, Appli
32	53.8	6.7	168	2	US-08-267-803B-4	Sequence 4, Appli
33	53.8	6.7	171	1	US-08-469-802B-5	Sequence 5, Appli
34	53.8	6.7	171	2	US-08-267-803B-5	Sequence 1, Appli
35	53.6	6.7	2338	2	US-08-425-069-1	Sequence 1, Appli
36	53.6	6.7	2338	2	US-08-317-844B-1	Sequence 1, Appli
37	53.4	6.6	1659	3	US-09-083-351-3	Sequence 3, Appli
38	53.4	6.6	1659	4	US-09-083-352-3	Sequence 3, Appli
39	53.4	6.6	3946	3	US-09-083-351-1	Sequence 1, Appli
40	53.4	6.6	3946	4	US-09-083-352-1	Sequence 1, Appli
41	53	6.6	7218	1	US-08-232-463-14	Sequence 14, Appl
42	52.6	6.5	51259	3	US-08-781-891-209	Sequence 209, App
43	52.2	6.5	203	4	US-09-043-303-7	Sequence 7, Appli
44	52.2	6.5	2830	2	US-09-010-928B-1	Sequence 1, Appli
45	52	6.5	165	4	US-09-043-303-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1

US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match		10.3%	Score 82.8	DB 4	Length 1926;
Best Local Similarity		51.6%	Pred. No. 3e+08;		
Matches 189;		Conservative 0;	Mismatches 177;	Indels 0;	Gaps 0;
Qy	386	AGGAGCTGGACAGAACTGTGTGAGAGGGAGGGCGCCCTCCAGAACTGTCAGCGCAGCT	445		
Db	519	AGGGGACAGCAGAGAGAGGGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	578		
Qy	446	TTGAGGAGAGAGAGCTTCCCTCCAGCTGGCCCTACGAGAGCGGCGCGGCTGTCAGGG	505		
Db	579	AGGAGGGGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	638		
Qy	506	ACGAGCTGGAGGGCGGAGCCCAAGCGGCAACAGCTCAAGCAGCGCTTCGCAAGAGA	565		
Db	639	AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	698		
Qy	566	GCCAGCGCGCCAGCAGGCTCTGCACCTTCAGCTTTCAGCAGGAGAGAGCGGC	625		
Db	699	AGGGGACAGCAGAGAGGGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	758		
Qy	626	AGCTCGGCGAGGAGCTCGAGAGCTCTATGAGAGGAGCAGGAGCTGTGAGAGCAGCTCA	685		
Db	759	AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	818		
Qy	686	GGTCTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	745		
Db	819	AGGGGACAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	878		

```
Qy 746 TGAGGC 751
Db 879 AGGGGC 884

RESULT 2
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 10.3%; Score 82.8; DB 3; Length 2580;
Best Local Similarity 51.6%; Pred. No. 3.1e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAAACAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTGCAGCGCAGCT 445
Db 902 AGGGCCAGGAGCAGGAGGAGGGCGAGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Qy 446 TTGAGGAGAGGAGCTTGCCTCCAGCTTGGCCCTACAGGAGGGCGGGCGGCTGCAGGG 505
Db 962 AGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1021
Qy 506 ACAGAGCTGGAGGGCCCGAGGCCCAAGCGGCAACAGCTCAAGCAGGCTCGCAGAGAAGA 565
Db 1022 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
Qy 566 GCCAGCGCGCAGCAGAGGTCTTGCACCTTGCAGGTACTGACGTTTCAGCAGGAGAGCGCGC 625
Db 1082 AGGGCCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1141
Qy 626 AGCTCCGGCAGAGCTTCAGAGACCTCATGAGGAGCAGGACCTGCTGAGAGCAGGAGCTCA 685
Db 1142 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
Qy 686 GGTCTTACGAGAGGGAGAGACCAAGCTTCCGGCCCGCGCTGGAGGAGACCCAGTGGGAGG 745

Db 1202 AGGGCAGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1261
Qy 746 TGAGGC 751
Db 1262 AGGGGC 1267

RESULT 3
US-09-359-081-2
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,081
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-359-081-2

Query Match 10.3%; Score 82.8; DB 4; Length 2580;
Best Local Similarity 51.6%; Pred. No. 3.1e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAAACAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTGCAGCGCAGCT 445
Db 902 AGGGCCAGGAGCAGGAGGAGGGCGAGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Qy 446 TTGAGGAGAGGAGCTTGCCTCCAGCTTGGCCCTACAGGAGGGCGGGCGGCTGCAGGG 505
Db 962 AGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1021
Qy 506 ACAGAGCTGGAGGGCCCGAGGCCCAAGCGGCAACAGCTCAAGCAGGCTCGCAGAGAAGA 565
Db 1022 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
Qy 566 GCCAGCGCGCAGCAGAGGTCTTGCACCTTGCAGGTACTGACGTTTCAGCAGGAGAGCGCGC 625
Db 1082 AGGGCCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1141
Qy 626 AGCTCCGGCAGAGCTTCAGAGACCTCATGAGGAGCAGGACCTGCTGAGAGCAGGAGCTCA 685
Db 1142 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
Qy 686 GGTCTTACGAGAGGGAGAGACCAAGCTTCCGGCCCGCGCTGGAGGAGACCCAGTGGGAGG 745
```

Db 1082 AGGGCAGGAGCAGGAGGGGCGAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1141
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGACCTGCTGAGACCAAGCTCA 685
Db 1142 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1201
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745
Db 1202 AGGGCAGGAGGGGCGAGGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 1261
Qy 746 TGAGGC 751
Db 1262 AGGGGC 1267

RESULT 4

US-09-130-114-1/c
; Sequence 1, Application US/09130114

; Patent No. 5976807

; GENERAL INFORMATION:

; APPLICANT: Horlick, Robert A.

; APPLICANT: Damaj, Bassam B.

; APPLICANT: Robbins, Alan K.

; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes

; FILE REFERENCE: 0867/1D903US1

; CURRENT APPLICATION NUMBER: US/09/130,114

; CURRENT FILING DATE: 1998-08-06

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 5452

; TYPE: DNA

; ORGANISM: VEBNA

US-09-130-114-1

Query Match 10.3%; Score 82.8; DB 2; Length 5452;

Best Local Similarity 51.6%; Pred. No. 3.4e-08;

Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAGCTGTGGAGAGGGGCGCCCTCCAGAGCTGAGCGCAGCT 445
Db 1903 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1844
Qy 446 TTGAGGAGAGGAGCTTCCTCCAGCTTGGCTTACGAGGAGCGCGCGCGCTGCAGGG 505
Db 1843 AGGAGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1784
Qy 506 AGGAGCTGGAGGGCCCGAGGCCAAAGCGCGCAACAGCTCAAGAGCGCTCGCAGAGA 565
Db 1783 AGGAGCAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAG 1724
Qy 566 GCCAGCGCGCAGCAGCTCTGCACCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 625
Db 1723 AGGGCAGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGG 1664
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGACCTGCTGAGACCAAGCTCA 685
Db 1663 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAG 1604
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745
Db 1603 AGGGCAGGAGGGGCGAGGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAG 1544
Qy 746 TGAGGC 751
Db 1543 AGGGGC 1538

RESULT 5

US-08-910-647-1

; Sequence 1, Application US/08910647

; Patent No. 6251433

; GENERAL INFORMATION:

; APPLICANT: Zuckermann et al.

; TITLE OF INVENTION: Compositions and Methods for

; TITLE OF INVENTION: Polynucleotide Delivery

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 94608-2916

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/910,647

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Fujita, Sharon M.

; REGISTRATION NUMBER: 38,459

; REFERENCE/DOCKET NUMBER: 1218.002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 923-2706

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9600 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-910-647-1

Query Match 10.3%; Score 82.8; DB 4; Length 9600;

Best Local Similarity 51.6%; Pred. No. 3.7e-08;

Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAGCTGTGGAGAGGGGCGCCCTCCAGAGCTGAGCGCAGCT 445
Db 948 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1007
Qy 446 TTGAGGAGAGGAGCTTCCTCCAGCTTGGCTTACGAGGAGCGCGCGCTGCAGGG 505
Db 1008 AGGAGGGGCGAGGAGCAGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1067
Qy 506 ACAGAGCTGGAGGGCCCGAGGCCAAAGCGCGCAACAGCTCAAGCAGGCGCTCGCAGAGA 565
Db 1068 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1127
Qy 566 GCCAGCGCGCAGCAGAGTCTGCACCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 625
Db 1128 AGGGCAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1187
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGACCTGCTGAGACCAAGCTCA 685
Db 1188 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAG 1247
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745
Db 1248 AGGGCAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 1307
Qy 746 TGAGGC 751
Db 1308 AGGGGC 1313

RESULT 6

US-09-620-925-1

; Sequence 1, Application US/09620925

Query Match	10.3%	Score 82.8	DB 4	Length 9600
Best Local Similarity	51.6%	Pred. No. 3.7e-08		
Matches 189	Conservative 0	Mismatches 177	Indels 0	Gaps 0
Qy	386	AGGAGCTGGGAACAGAACTGTTGCAGAGGGAGGGCGCCCTCCAGAACTGCAGCGCAGCT	445	
Db	948	AGGGGCAGGAGCAGGAGAGGGGCAGAGAGGGCAGGAGCAGGAGGGGCAGGAGGGGC	1007	
Qy	446	TTGAGGAGAAGGAGCTTTGCCCTCCAGCTTGCGCTTACGAGGAGCGGCCCGCGCTGTCAGGG	505	
Db	1008	AGGAGGGCGAGGACAGGAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGGGC	1067	
Qy	506	ACGAGCTGGAGGGGCCGAGAGCCCAAGGGCGCACAGCTCAAGCAGGCCTCGCAGAAGA	565	
Db	1068	AGGAGCAGGAGGGGCAGGAGCAGGAGGGGGCAGGAGGGGCAGGAGGGGCAGGAGCAGG	1127	
Qy	566	GCCAGCGCGCAGCAGGTCTCTGCACCTGCAGGTACTGTCAGCTTTCAGCAGGAGAAGCGGC	625	
Db	1128	AGGGGCAGGAGCAGGAGGGGCAGAGCAGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGGC	1187	
Qy	626	AGTTCGCGCAGGAGCTCGAGAGCCTCATGAGGAGCAGGACCTCTCTGGAGACCAGTCTCA	685	
Db	1188	AGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGG	1247	
Qy	686	GGTCCTACGAGAGGGAGAAGACCAGCTTCGSCCCCGCGCTTGAGGAGAGCCCGATGGGAGG	745	
Db	1248	AGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGGC	1307	
Qy	746	TGAGGC	751	
Db	1308	AGGGGC	1313	

	Query Match	10.3%	Score 82.8	DB 1	Length 10596
	Best Local Similarity	51.6%	Pred. No. 3.8e-08		
	Matches 189	Conservative 0	Mismatches 177	Indels 0	Gaps 0
Qy	386	AGGAGCTGGACAGACGCTGTTGGAGGAGGAGGGCCCTCCAGAGAGCTTCGACGGCAGCT	445		
Db	2482	AGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGACGAGGAGGAGGGGCAGGAGGGGC	2541		
Qy	446	TTGAGGAGAAGGAGCTTGCTCCAGCTGTGGCTTACGAGGAGCGCGCGCGCTGCAGGG	505		
Db	2542	AGAGGGCAGGACGAGGAGGAGGGCAGGAGCAGAGGGCAGGAGGGGCAGGAGGGGC	2601		
Qy	506	ACGAGCTGGAGGGCCCGGAGCCCAAGCGCGCAACAGTCTAAGCAGGCTTCGACGAGA	565		
Db	2602	AGGAGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGCAGG	2661		
Qy	566	CCAGCGCGCGCAGCAGGTCCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGNAGCGGC	625		
Db	2662	AGGGGCAGGAGCAGGAGGGGCAGGACGAGAGGGGCAGGACGAGGAGGAGGGGCAGGAGGGGC	2721		
Qy	626	AGCTCCGGCAGGAGCTCGAGAGCTCATGTAAGGAGCAGGACCTGCTGGACACCAAGCTCA	685		
Db	2722	AGNGCAGAGAGGGCAGGAGGGGCAGGAGCAGAGAGGGGCAGGAGGGGCAGGACGAGGAGG	2781		
Qy	686	GGTCCTTACGAGAGGGAGAGAAGCCAGCTTCGCGCCCGCGCTTGGAGGAGAGACCCAGTGGGAGG	745		
Db	2782	AGGGCAGGAGGGGCAGGACGAGGAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGGGCAGGAGGGGCAGG	2841		


```
Qy 746 TGAGGC 751
Db 2842 AGGGGC 2847

RESULT 8
US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregler, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-15

Query Match 10.3%; Score 82.8; DB 1; Length 10596;
Best Local Similarity 51.6%; Pred. No. 3.8e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAGCTTGTTCAGAGAGGGGGCCCTCCAGAGCTGCAGCGCAGCT 445
Db 2482 AGGGGACAGCAGGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2541
Qy 446 TTGAGGAGAGAGCTTGCCTCCAGCTGGCTTACGAGAGCGGCGCGCGCTGCAGGG 505
Db 2542 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2601
Qy 506 ACAGCTGGAGGGCCCGAGGCGCAAGAGCGGCAACAGCTCAAGCAGCGCTTCGAGAGA 565
Db 2602 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2661
Qy 566 GCCAGCGGCGCAGCAGCTCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGAGCGGC 625
Db 2662 AGGGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2721
Qy 626 AGCTCCGCGCAGGAGCTTCAGAGCTTCATGAAAGGAGCAGGAGCTTGTGTGAGACCAAGCTCA 685
Db 2722 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2781
Qy 686 GGTCTTACGAGAGGAGAGAACAGCTTTCGCGCCCGCGCTCGAGAGGAGACCAGTGGGAGG 745
Db 2842 AGGGGC 2847

US-08-087-783A-15
; Sequence 15, Application US/08087783A
; Patent No. 5547856
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-087-783A-15

Query Match 10.3%; Score 82.8; DB 1; Length 10596;
Best Local Similarity 51.6%; Pred. No. 3.8e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAGCTTGTTCAGAGAGGGGGCCCTCCAGAGCTGCAGCGCAGCT 445
Db 2482 AGGGGACAGCAGGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2541
Qy 446 TTGAGGAGAGAGCTTGCCTCCAGCTGGCTTACGAGAGCGGCGCGCGCTGCAGGG 505
Db 2542 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2601
Qy 506 ACAGCTGGAGGGCCCGAGGCGCAAGAGCGGCAACAGCTCAAGCAGCGCTTCGAGAGA 565
Db 2602 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2661
Qy 566 GCCAGCGGCGCAGCAGCTCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGAGCGGC 625
Db 2662 AGGGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2721
Qy 626 AGCTCCGCGCAGGAGCTTCAGAGCTTCATGAAAGGAGCAGGAGCTTGTGTGAGACCAAGCTCA 685
Db 2722 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2781
Qy 686 GGTCTTACGAGAGGAGAGAACAGCTTTCGCGCCCGCGCTCGAGAGGAGACCAGTGGGAGG 745
Db 2842 AGGGGC 2847
```

QY	626	AGTCCGGCAGGAGCTCGAGAGCTCTAAGAGGACAGGACCTCTCTGGAGACCAAGTCA	685
Db	2722	AGGAGCAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGG	2781
QY	686	GGTCTTACGAGAGGGAGAAGACCAGCTTCGGCCCCCGCTGGAGGAGACCCAGTGGGAGG	745
Db	2782	AGGGCAGGAGGGGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGGGCAGG	2841
QY	746	TGAGGC	751
Db	2842	AGGGGC	2847

RESULT 10
US-08-194-088B-15
; Sequence 15, Application US/08194088B
; Patent No. 5580963
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,088B
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallegos, R. Thomas
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: 755D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-2614
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-194-088B-15

Qy	566	GCACAGCGCGGACGACAGGTCTCTCACCTCGAGGTACTGTGAGCTTTCAGCAGGAGAAAGCGGC	625
Db	2662	AGCGGACGAGGACGAGGAGGGGCGAGGACGAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGC	2721
Qy	626	AGCTCCGGCAGGAGCTCGAGAGCGCTCATGAAGAGGACGAGACCTGCTGGAGACCAAGCTCA	685
Db	2722	AGGAGCAGGAGGGCGCAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGCGAGGACGAGGAGG	2781
Qy	686	GGTCTTACGAGAGGGGAGAGACCCAGCTTCGCGCCCCCGCTGGAGGAGAGACCCAGTGGGAGG	745
Db	2782	AGGGGACGAGGGGCGCAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG	2841
Qy	746	TGAGGC 751	
Db	2842	AGGGGC 2847	
RESULT 11			
US-08-194-087-15			
; Sequence 15, Application US/08194087			
; Patent No. 5879910			
; GENERAL INFORMATION:			
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.			
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Genentech, Inc.			
; STREET: 460 Point San Bruno Blvd			
; CITY: South San Francisco			
; STATE: California			
; COUNTRY: USA			
; ZIP: 94080			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: patin (Genentech)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/194,087			
; FILING DATE: 18-MAY-1992			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Dreger, Ginger R.			
; REGISTRATION NUMBER: 33,055			
; REFERENCE/DOCKET NUMBER: 779			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415/225-3216			
; TELEFAX: 415/952-9881			
; TELEX: 910/371-7168			
; INFORMATION FOR SEQ ID NO: 15:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 10596 bases			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-194-087-15			
Query Match 10.3%; Score 82.8; DB 2; Length 10596;			
Best Local Similarity 51.6%; Prsd. No. 3.8e-08;			
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;			
Qy	386	AGAGCTGGAAACAGAGCTTTTGGAGAGGAGGGGCGCTTCAGAGCTTCAGGCGAGCT	445
Db	2482	AGGGGACGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGC	2541
Qy	446	TTGAGGAGAGGAGCTTTCCTCCAGCTGCGCTACGAGGAGCGCGCGCGCTGACAGG	505
Db	2542	AGGAGGGGCGAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGC	2601
Qy	506	ACGAGCTGGAGGGCGCGGAGCCCAAGGGCGCAACAGCTTCAAGCAGGGCCTCGCAGAGA	565

Qy	565	AGCCAGCGCGCGACGAGTCTCTGCACTG	CAGAGTACTG	CAGCTT	CAGCAGGAGAACGCG	634
Db	2179	GATCAGCAGCAGCAGCAGGATGAGCAGCAG	CAGCAGGATCAGCAGCAGCAGCAGCAGGATGAG			2238
Qy	625	CAGCTCCGGCAGGAGCTCGAGAGCCTCAT	GAGGAGCAGGACCTGCTGGAGACCAAGCTC			684
Db	2239	CAGCAGCAGGAGTGAACAGGAGCAGCAGG	AGGAGCAGCAGCAGCAGGAGCAGGAGCAGGAG			2298
Qy	685	AGGTCCTACGAGGAGGAGAACAGCTTTCGG	CCCCCGCTGGAGGAGACCCAGTGGGAG			744
Db	2299	CAGGAGTTAGAGGAGCAGGAGCAGGATCAG	GAGCAGGAGGATAGAGGATAGAGGAGCAG			2358
Qy	745	GTG	747			
Db	2359	GAG	2361			

Search completed: June 15, 2003, 03:45:46
Job time : 49.1752 sec8

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 02:18:56 ; Search time 117.039 Seconds
(without alignments)
9947.427 Million cell updates/sec

Title: US-09-513-888c-1_COPY_1707_2510

Perfect score: 804

Sequence: 1 gccatctgactctcccc.....tttggcgccagtaccacct 804

Scoring table: IDENTITY NUC
Gapop 10^0, Gapext 1.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/PCT_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	802.4	99.8	2411	9	US-09-764-904-89
2	802.4	99.8	2411	9	US-10-091-548-89
3	802.4	99.8	2411	9	US-10-074-095-1133
4	802.4	99.8	2411	10	US-09-764-860-1133
5	463.4	57.6	560	9	US-09-764-904-17
6	463.4	57.6	560	9	US-10-091-548-17
7	463.4	57.6	560	9	US-10-074-095-269
8	463.4	57.6	560	10	US-09-764-860-269
9	151	18.8	5257	9	US-10-171-581-63
10	112.4	14.0	561	10	US-09-917-800A-194
11	67	8.3	1617	10	US-09-925-300-60
12	67	8.3	2108	10	US-09-962-832-225
13	57.8	7.2	8459	9	US-09-563-728A-31
14	57.8	7.2	8459	10	US-09-817-913-8
15	57.8	7.2	8459	10	US-09-817-538-8
16	55.6	6.9	14800	10	US-09-954-456-1601
17	55.4	6.9	4366	12	US-10-044-090-668
18	55	6.8	596	10	US-09-864-761-8648
19	55	6.8	2145	9	US-10-153-668-125

20	55	6.8	2145	9	US-10-153-668-127	Sequence 127, App
21	54.4	6.8	2313	9	US-09-854-133-157	Sequence 157, App
22	54.4	6.8	2313	10	US-09-738-973-157	Sequence 157, App
23	54.4	6.8	2314	9	US-09-764-868-12	Sequence 12, Appl
24	54.4	6.8	2314	10	US-09-778-927A-16	Sequence 16, Appl
25	53.8	6.7	4282	10	US-09-864-761-25366	Sequence 25366, A
26	53.8	6.7	2785	10	US-09-925-302-122	Sequence 122, App
27	53.6	6.7	6604	10	US-09-880-107-1748	Sequence 1748, App
28	53.4	6.6	1662	9	US-09-292-862-1	Sequence 1, Appli
29	53.2	6.6	1649	9	US-09-789-390-62	Sequence 62, Appl
30	53.2	6.6	1787	9	US-09-789-390-66	Sequence 66, Appl
31	53.2	6.6	3822	9	US-09-789-390-10	Sequence 10, Appl
32	53.2	6.6	3971	9	US-10-270-333-59	Sequence 59, Appl
33	53.2	6.6	3999	9	US-09-789-390-6	Sequence 6, Appli
34	53.2	6.6	4133	9	US-09-291-417-11	Sequence 11, Appl
c	53.2	6.6	7503	9	US-10-270-333-58	Sequence 58, Appl
36	53	6.6	422	9	US-09-854-133-337	Sequence 337, App
37	53	6.6	422	10	US-09-738-973-337	Sequence 337, App
38	52.8	6.6	22585	9	US-09-764-891-6987	Sequence 6987, App
39	52.6	6.5	2614	9	US-09-822-846-491	Sequence 491, App
40	52.6	6.5	4033	9	US-10-029-115-5	Sequence 5, Appli
41	52.2	6.5	1852	10	US-09-969-852-4	Sequence 4, Appli
42	51.2	6.4	1827	9	US-09-974-298-174	Sequence 174, App
43	51.2	6.4	1827	10	US-09-919-172-48	Sequence 48, Appl
44	51.2	6.4	6457	10	US-09-880-107-3389	Sequence 3389, App
45	51	6.3	520	9	US-10-184-644-332	Sequence 332, App

ALIGNMENTS

RESULT 1

US-09-764-904-89
; Sequence 89, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764, 904
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-89

Query Match	99.8%	Score 802.4;	DB 9;	Length 2411;
Best Local Similarity	99.9%	Pred. No. 6.8e-187;		
Matches 803;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC	60	
Db	1367	GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC	1426	
Qy	61	AAGCCCAAGGAGGAGGAGCTGAAGCCTGGCTGTGTCTTGGGGGCGCTGTAGACTCCGGC	120	
Db	1427	AAGCCCAAGGAGGAGGAGCTGAAGCCTGGCTGTGTCTTGGGGGCGCTGTAGACTCCGGC	1486	
Qy	121	CGGAATCTCAATGTCTCCAGCTGCCACACACAGCAGCAGCAGCTACCACTGACCCCG	180	
Db	1487	CGGAATCTCAATGTCTCCAGCTGCCACACACAGCAGCAGCAGCTACCACTGACCCCG	1546	
Qy	181	CTGTGTCAACCCGTTGGGACCCCAAGCCGTTTTGGGGGCTCCGCCCAACATCACCAG	240	
Db	1547	CTGTGTCAACCCGTTGGGACCCCAAGCCGTTTTGGGGGCTCCGCCCAACATCACCAG	1606	
Qy	241	GGCATCTGCTCTCCAGGACAGCAATATAGCTGAAGGCTCTGTCTTCTTCGACCGA	300	
Db	1607	GGCATCTGCTCTCCAGGACAGCAATATAGCTGAAGGCTCTGTCTTCTTCGACCGA	1666	

QY 301 GGTAGCAAGCTGGGCACTCGAACAAGGACAGACAAAGGCCCCCTCGTGTGTCGCTCCCCC 360
DB 1667 GGTAGCAAGCTGGGCACTCGAACAAGGACAGACAAAGGCCCCCTCGTGTGTCGCTCCCCC 1726
QY 361 ATCTCCACGACGAGTGCAGCATCAGAGAGCTGGAACAAGAGCTGTGTTGGAGAGGAGGGC 420
DB 1727 ATCTCCACGACGAGTGCAGCATCAGAGAGCTGGAACAAGAGCTGTGTTGGAGAGGAGGGC 1786
QY 421 GCCCTCCAGAAGCTGCAGCGAGCTTTGAGAGAGAGCTTGCTCCAGCCTGGCCTTAC 480
DB 1787 GCCCTCCAGAAGCTGCAGCGAGCTTTGAGAGAGAGCTTGCTCCAGCCTGGCCTTAC 1846
QY 481 GAGGAGCGGCGCGGCGCTGCAGGACAGCTGAGGCGCGGAGCCCAAGGCGGCAAC 540
DB 1847 GAGGAGCGGCGCGGCGCTGCAGGACAGCTGAGGCGCGGAGCCCAAGGCGGCAAC 1906
QY 541 AAGCTCAAGCAGGCTCGCAGAAGAGCCAGCGCGCAGCAGAGCTTCTGCACCTGCAGGTA 600
DB 1907 AAGCTCAAGCAGGCTCGCAGAAGAGCCAGCGCGCAGCAGAGCTTCTGCACCTGCAGGTA 1966
QY 601 CTGACGCTTCAGCAGGAGAGCGGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAAGGAG 660
DB 1967 CTGACGCTTCAGCAGGAGAGCGGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAAGGAG 2026
QY 661 CAGGACCTGCTGGAGACCAAGCTCAGTCTTACGAGGAGGAGAGACAGCTTCGGCCCC 720
DB 2027 CAGGACCTGCTGGAGACCAAGCTCAGTCTTACGAGGAGGAGAGACAGCTTCGGCCCC 2086
QY 721 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTGGGTGGTCA 780
DB 2087 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTGGGTGGTCA 2146
QY 781 CGGTTTGGCCCAAGTACCCCCCT 804
DB 2147 CGGTTTGGCCCAAGTACCCCCCT 2170

RESULT 2

US-10-091-548-89
; Sequence 89, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-89

Query Match. 99.8%; Score 802.4; DB 9; Length 2411;
Best Local Similarity 99.9%; Pred. No. 6.8e-187;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCATCTGCACTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
DB 1367 GCCATCTGCACTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 1426
QY 61 AAGCCCAAGGACGAGAGCTGAAGCTGGGCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 120
DB 1427 AAGCCCAAGGACGAGAGCTGAAGCTGGGCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 1486
QY 121 CGGAATCTCATGTCCAGCTGCGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 180
DB 1487 CGGAATCTCATGTCCAGCTGCGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 1546

QY 181 CTGCTCACACCCGTGGGACCCCAAGCGCTTTGGGGCTCCCGCCACACATCACCCAG 240
DB 1547 CTGCTCACACCCGTGGGACCCCAAGCGCTTTGGGGCTCCCGCCACACATCACCCAG 1606
QY 241 GGATCTCTCTCCAGGACAGCAACATGATGAGCTCTGAAGGCTCTGTCTTCTCCGACGGA 300
DB 1607 GGATCTCTCTCCAGGACAGCAACATGATGAGCTCTGAAGGCTCTGTCTTCTCCGACGGA 1656
QY 301 GGTAGCAAGCTGGGCACTCGAACAAGGACAGACAAAGGCCCCCTCGTGTGTCGCTCCCCC 360
DB 1667 GGTAGCAAGCTGGGCACTCGAACAAGGACAGACAAAGGCCCCCTCGTGTGTCGCTCCCCC 1726
QY 361 ATCTCCACGACGAGTGCAGCATCAGAGAGCTGGAACAAGAGCTGTGTTGGAGAGGAGGGC 420
DB 1727 ATCTCCACGACGAGTGCAGCATCAGAGAGCTGGAACAAGAGCTGTGTTGGAGAGGAGGGC 1786
QY 421 GCCCTCCAGAAGCTGCAGCGAGCTTTGAGAGAGAGCTTGCTCCAGCCTGGCCTTAC 480
DB 1787 GCCCTCCAGAAGCTGCAGCGAGCTTTGAGAGAGAGCTTGCTCCAGCCTGGCCTTAC 1846
QY 481 GAGGAGCGGCGCGGCGCTGCAGGACAGCTGAGGCGCGGAGCCCAAGGCGGCAAC 540
DB 1847 GAGGAGCGGCGCGGCGCTGCAGGACAGCTGAGGCGCGGAGCCCAAGGCGGCAAC 1906
QY 541 AAGCTCAAGCAGGCTCGCAGAAGAGCCAGCGCGCAGCAGAGCTTCTGCACCTGCAGGTA 600
DB 1907 AAGCTCAAGCAGGCTCGCAGAAGAGCCAGCGCGCAGCAGAGCTTCTGCACCTGCAGGTA 1966
QY 601 CTGACGCTTCAGCAGGAGAGCGGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAAGGAG 660
DB 1967 CTGACGCTTCAGCAGGAGAGCGGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAAGGAG 2026
QY 661 CAGGACCTGCTGGAGACCAAGCTCAGTCTTACGAGGAGGAGAGACAGCTTCGGCCCC 720
DB 2027 CAGGACCTGCTGGAGACCAAGCTCAGTCTTACGAGGAGGAGAGACAGCTTCGGCCCC 2086
QY 721 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTGGGTGGTCA 780
DB 2087 CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGGTTGGGTGGTCA 2146
QY 781 CGGTTTGGCCCAAGTACCCCCCT 804
DB 2147 CGGTTTGGCCCAAGTACCCCCCT 2170

RESULT 3

US-10-074-095-1133
; Sequence 1133, Application US/10074095
; Publication No. US2003007704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC08C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08

```
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      99.8%; Score 802.4; DB 9; Length 2411;
Best Local Similarity 99.9%; Pred. No. 6.8e-187;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 1367 GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 1426

Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTAGACTCCGGC 120
Db 1427 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTAGACTCCGGC 1486

Qy 121 CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCCG 180
Db 1487 CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCCG 1546

Qy 181 CTGCTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCGGCCAACATACCCAG 240
Db 1547 CTGCTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCGGCCAACATACCCAG 1606

Qy 241 GGCATCTGCTCCAGGACAGCAACATGATGAGCTGTGCTGGGGCGCTGTAGACTCCGGC 300
Db 1607 GGCATCTGCTCCAGGACAGCAACATGATGAGCTGTGCTGGGGCGCTGTAGACTCCGGC 1666

Qy 301 GGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGGCCCTCTGTGTCTCGCTCCGCC 360
Db 1667 GGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGGCCCTCTGTGTCTCGCTCCGCC 1726

Qy 361 ATCTCCAGGACAGTGAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGGC 420
Db 1727 ATCTCCAGGACAGTGAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGGC 1786

Qy 421 GCCCTCCAGAGCTGCAGCGAGCTTTCAGGAGAGGAGCTTGCTCCAGCTGCCTTAC 480
Db 1787 GCCCTCCAGAGCTGCAGCGAGCTTTCAGGAGAGGAGCTTGCTCCAGCTGCCTTAC 1846

Qy 481 GAGGAGCGCGCGCGCTGACAGAGAGCTGTGGAGGGCCCGGAGCCCAAGGGCGCAAC 540
Db 1847 GAGGAGCGCGCGCGCTGACAGAGAGCTGTGGAGGGCCCGGAGCCCAAGGGCGCAAC 1906

Qy 541 AAGCTCAAGCAGGCTTCAGAGAGAGCCAGCGCGCGCAGAGGCTTCTGCACCTGCAGTA 600
Db 1907 AAGCTCAAGCAGGCTTCAGAGAGAGCCAGCGCGCGCAGAGGCTTCTGCACCTGCAGTA 1966

601 CTGCAGCTTCAGCAGGAGAAAGCGCAGCTCGGCGAGAGCTCGAGAGCTTCATGAAGGAG 660
1967 CTGCAGCTTCAGCAGGAGAAAGCGCAGCTCGGCGAGAGCTCGAGAGCTTCATGAAGGAG 2026
661 CAGGACCTGTCTGGAGACCAAGCTCAGCTCCTTACGAGAGGAGAAAGACAGCTTCGGCCCC 720
2027 CAGGACCTGTCTGGAGACCAAGCTCAGCTCCTTACGAGAGGAGAAAGACAGCTTCGGCCCC 2086
721 GCGCTGAGGAGAGACCCAGTGGAGGTGAGGCCACACAGGCTCATGGTTTGGGTGTCTCA 780
2087 GCGCTGAGGAGAGACCCAGTGGAGGTGAGGCCACACAGGCTCATGGTTTGGGTGTCTCA 2146
781 GCGGTTTGGCGCCAGTACCCCT 804
2147 GCGGTTTGGCGCCAGTACCCCT 2170

RESULT 4
US-09-764-860-1133
; Sequence 1133, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1133
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1133

Query Match      99.8%; Score 802.4; DB 10; Length 2411;
Best Local Similarity 99.9%; Pred. No. 6.8e-187;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 1367 GCCATCTGCACTCTCTCCCGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 1426

Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTAGACTCCGGC 120
Db 1427 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTAGACTCCGGC 1486

Qy 121 CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCCG 180
Db 1487 CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGGACCCG 1546

Qy 181 CTGCTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCGGCCAACATACCCAG 240
Db 1547 CTGCTCAACCCGTGGGACCCACAAAGCCGTTTGGGGGCTCGGCCAACATACCCAG 1606

Qy 241 GGCATCTGCTCCAGGACAGCAACATGATGAGCTGTGCTGGGGCGCTGTAGACTCCGGC 300
Db 1607 GGCATCTGCTCCAGGACAGCAACATGATGAGCTGTGCTGGGGCGCTGTAGACTCCGGC 1666

Qy 301 GGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGGCCCTCTGTGTCTCGCTCCGCC 360
Db 1667 GGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGGCCCTCTGTGTCTCGCTCCGCC 1726

Qy 361 ATCTCCAGGACAGTGAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGGC 420
Db 1727 ATCTCCAGGACAGTGAGCATCCAGAGCTGGAACAGAGCTGTGGAGAGGAGGGC 1786

Qy 421 GCCCTCCAGAGCTGCAGCGAGCTTTCAGGAGAGGAGCTTGCTCCAGCTGCCTTAC 480
Db 1787 GCCCTCCAGAGCTGCAGCGAGCTTTCAGGAGAGGAGCTTGCTCCAGCTGCCTTAC 1846

Qy 481 GAGGAGCGCGCGCGCTGACAGAGAGCTGTGGAGGGCCCGGAGCCCAAGGGCGCAAC 540
Db 1847 GAGGAGCGCGCGCGCTGACAGAGAGCTGTGGAGGGCCCGGAGCCCAAGGGCGCAAC 1906

Qy 541 AAGCTCAAGCAGGCTTCAGAGAGAGCCAGCGCGCGCAGAGGCTTCTGCACCTGCAGTA 600
Db 1907 AAGCTCAAGCAGGCTTCAGAGAGAGCCAGCGCGCGCAGAGGCTTCTGCACCTGCAGTA 1966
```

Db 1847 GAGGAGCGCGCGCGCTGAGGAGCAGCTGAGGCGCGGAGCCCAAGGGGGCAAC 1906
Qy 541 AAGCTCAAGCAGGCTCGCAGAGAGCAGCGCGCAGCAGGCTCTGCACCTGCAGTA 600
Db 1907 AAGCTCAAGCAGGCTCGCAGAGAGCAGCGCGCGCAGCAGGCTCTGCACCTGCAGTA 1966
Qy 601 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAG 660
Db 1967 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAG 2026
Qy 661 CAGGACCTCTGAGAGACCAAGCTCAGTCTTACGAGAGGAGAGACAGCTTCGGCCCC 720
Db 2027 CAGGACCTCTGAGAGACCAAGCTCAGTCTTACGAGAGGAGAGACAGCTTCGGCCCC 2086
Qy 721 CGCTGAGGAGAGCAGCTGAGGAGTGAGGAGCAGCAGCGCTCATGGTTGGGTGTCA 780
Db 2087 CGCTGAGGAGAGCAGCTGAGGAGTGAGGAGCAGCAGCGCTCATGGTTGGGTGTCA 2146
Qy 781 GCGGTTTGGCGCCAGTACCCCCCT 804
Db 2147 GCGGTTTGGCGCCAGTACCCCCCT 2170

RESULT 5

US-09-764-904-17
; Sequence 17, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-17

Query Match 57.6%; Score 463.4; DB 9; Length 560;
Best Local Similarity 99.8%; Pred. No. 3.6e-104;
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCCATCTGCACTCTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db 96 GCCATCTGCACTCTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 155
Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 120
Db 156 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 215
Qy 121 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCTACCACTGCAGCTCCGCC 180
Db 216 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCTACCACTGCAGCTCCGCC 275
Qy 181 CTGCTCACACCCGTGGAGCCACCAAGCGTGTGCTCTGGGGCGCTGTGAGACTCCGCC 240
Db 276 CTGCTCACACCCGTGGAGCCACCAAGCGTGTGCTCTGGGGCGCTGTGAGACTCCGCC 335
Qy 241 GGCATCTGCTCTCCAGGACAGCAACATGATGAGCTGGAAGGCTCTGCTCTCCGACGGA 300
Db 336 GGCATCTGCTCTCCAGGACAGCAACATGATGAGCTGGAAGGCTCTGCTCTCCGACGGA 395
Qy 301 GGTAGCAGCTGGGCACTCGAACAAGCAGACAGAGGCGCCCTCGTGTGCTCCGCC 360
Db 396 GGTAGCAGCTGGGCACTCGAACAAGCAGACAGAGGCGCCCTCGTGTGCTCCGCC 455
Qy 361 ATCTCCAGCAGCTGAGTGCAGATCCAGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 420
Db 456 ATCTCCAGCAGCTGAGTGCAGATCCAGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 515

RESULT 7

US-10-091-548-17
; Sequence 17, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-17

RESULT 7

US-10-074-095-269
; Sequence 269, Application US/10074095
; Publication No. US2003007704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14

;; PRIOR APPLICATION NUMBER: 09/764,860
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 60/179,065
;; PRIOR FILING DATE: 2000-01-31
;; PRIOR APPLICATION NUMBER: 60/180,628
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: 60/214,886
;; PRIOR FILING DATE: 2000-06-28
;; PRIOR APPLICATION NUMBER: 60/217,487
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,758
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,963
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/217,496
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,447
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/218,290
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/225,757
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/226,868
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/216,647
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,267
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/216,880
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,270
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/251,869
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/235,834
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/234,274
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/234,223
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/228,924
;; PRIOR FILING DATE: 2000-08-30
;; PRIOR APPLICATION NUMBER: 60/224,518
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,369
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/224,519
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,964
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/241,809
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17

;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08

Query Match 57.6%; Score 463.4; DB 9; Length 560;
Best Local Similarity 99.8%; Pred. No. 3.6e-104;
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 155

Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTGCTCTGGGGCGCTGTCTAGACTCCGGC 120
Db AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTGCTCTGGGGCGCTGTCTAGACTCCGGC 215

Qy 121 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGAGCTACAGCTGAGCCCG 180
Db CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGAGCTACAGCTGAGCCCG 275

Qy 181 CTGGTTCACCCGTTGGGACCCACAGCCGTTTGGGGCTCGGCCCAACATCACCAG 240
Db CTGGTTCACCCGTTGGGACCCACAGCCGTTTGGGGCTCGGCCCAACATCACCAG 335

Qy 241 GGCATCGTCTCCAGGACAGCAACATGATAGCCTGAAGGCTGTGCTTCTCCGACGGA 300
Db GGCATCGTCTCCAGGACAGCAACATGATAGCCTGAAGGCTGTGCTTCTCCGACGGA 395

Qy 361 ATCTCCACGAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 420
Db ATCTCCACGAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 515

Qy 421 GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 465
Db GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 560

RESULT 9
US-10-171-581-63
; Sequence 63, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue

Db 396 GGTAGCAAGCTGGCCACTCGAACAGGCACACAAGGGCCCCCTCGTGTGTCGGCTCCCC 455
Qy 361 ATCTCCACGAGCTGGCCACTCGAACAGGCACACAAGGGCCCCCTCGTGTGTCGGCTCCCC 420
Db 456 ATCTCCACGAGCTGGCCACTCGAACAGGCACACAAGGGCCCCCTCGTGTGTCGGCTCCCC 515
Qy 421 GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 465
Db 516 GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 560

RESULT 8
US-09-764-860-269
; Sequence 269, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-269

Query Match 57.6%; Score 463.4; DB 10; Length 560;
Best Local Similarity 99.8%; Pred. No. 3.6e-104;
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60
Db GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 155

Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTGCTCTGGGGCGCTGTCTAGACTCCGGC 120
Db AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTGCTCTGGGGCGCTGTCTAGACTCCGGC 215

Qy 121 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGAGCTACAGCTGAGCCCG 180
Db CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGAGCTACAGCTGAGCCCG 275

Qy 181 CTGGTTCACCCGTTGGGACCCACAGCCGTTTGGGGCTCGGCCCAACATCACCAG 240
Db CTGGTTCACCCGTTGGGACCCACAGCCGTTTGGGGCTCGGCCCAACATCACCAG 335

Qy 241 GGCATCGTCTCCAGGACAGCAACATGATAGCCTGAAGGCTGTGCTTCTCCGACGGA 300
Db GGCATCGTCTCCAGGACAGCAACATGATAGCCTGAAGGCTGTGCTTCTCCGACGGA 395

Qy 301 GGTAGCAAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 360
Db GGTAGCAAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 455

Qy 361 ATCTCCACGAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 420
Db ATCTCCACGAGCTGGCCCAACAGCAGCAGAGCTGAGAGCTGTTGGAGGAGGAGGC 515

Qy 421 GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 465
Db GGCCTCCAGAGCTGCAGCCAGCTTTGAGGAGAGGAGCTTGCC 560

```

; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 194
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA875041
; US-09-917-800A-194

Query Match 14.0%; Score 112.4; DB 10; Length 561;
Best Local Similarity 77.0%; Pred. No. 1.8e-18; Mismatches 41; Indels 0; Gaps 0;
Matches 137; Conservative 0;

QY 367 ACGACGAGTGCAGCATCCAGGAGCTTGGAAACAGAAAGCTGTTGAGAGGGAGGGCGCCCTC 426
Db |||||
561 ACGGATGAGTGCACCATTCAGGAGCTGGAGCAGAAAGCTGCTGCAGCGGGAGACTGCACCTA 502
QY 427 CAGAAGCTCAGCGCAGCTTTCAGGAGAGAGGAGCTTGCCTCCAGCCCTGCGCTACGAGGAG 486
Db |||||
501 CAGAAGCTCAGCGCAGCTTTCGATGAGAGAGGAGTTTGCCTCTGCGCAGACCTTCGAGGAG 442
QY 487 CGGCCCGCGCGCTGTCAGGAGCAGAGCTGGAGGGCCCGGAGGCCAAAGCGCAACAGC 544
Db |||||
441 CGGCCACGCGGAGCTAGACGAACTGGAGTGCCTTGSAACTTAAGACCAAGCTGAAGC 384

RESULT 11
US-09-925-300-60
; Sequence 60, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 60
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1590)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1592)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1595)
; OTHER INFORMATION: n equals a,t,g, or c

```

QY	686	GGTCTTACGAGAGGAGAACACCGCTTCGGCCCCCGCGTGTGAGGAG	732
DB	1270	AGCAGCAGGAGGCGAGGTGGAGCACCTGGAGCAGCAGGTGGGGCAG	1316
RESULT 13			
US-09-563-728A-31			
; Sequence 31, Application US/09563728A			
; Publication No. US20030078216A1			
; GENERAL INFORMATION:			
; APPLICANT: MacLeod, Alan R			
; APPLICANT: Li, Zoumei			
; APPLICANT: Besterman, Jeffrey M			
; TITLE OF INVENTION: Inhibition of Histone Deacetylase			
; FILE REFERENCE: 106101.229			
; CURRENT APPLICATION NUMBER: US/09/563,728A			
; CURRENT FILING DATE: 2000-05-03			
; PRIOR APPLICATION NUMBER: 60/132,287			
; PRIOR FILING DATE: 1999-05-03			
; NUMBER OF SEQ ID NOS: 36			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 31			
; LENGTH: 8459			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-563-728A-31			
Query Match 7.2%; Score 57.8; DB 9; Length 8459;			
Best Local Similarity 48.3%; Pred. No. 4.3e-05;			
Matches 161; Conservative 0; Mismatches 172; Indels 0; Gaps 0;			
QY	401	AGCTGTTGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGC	460
DB	980	AGCGGCGCCCTGCGGGAGCAGCAGCTGCAGCAGGAGCTCTCTGGCGCTCAAGCAGAAGCAGC	1039
QY	461	TTGCTTCCAGCTTGGCGCTACGAGGAGCGCGCGCGCGCTTGCAGGGAGCAGACTGAGGGGCC	520
DB	1040	AGATCCAGAGGCGAGATCTCATCTGCTAGTTTCCAGAGGCGAGCACGAGCAGCTCTCCCGC	1099
QY	521	CGGAGGCCAACAGGCGCAACAGCTCAAGCAGCGCTTCGAGAGAGAGCCAGCGCGCGCAGC	580
DB	1100	AGCAGAGGCGGAGCTCCACGAGCACATCAAGCAATAACAGGAGATGCTGGCCATGAAGC	1159
QY	581	AGGTCTTCGACCTTCGAGGTACTGCGCTTTCAGCAGGAGAGCGGCGAGCTCCGCGCAGGAGC	640
DB	1160	ACCAGCAGGAGCTGCTGGAACACCCAGCGGAGCTGGAGGGCACCCGCCAGGAGCAGGAGC	1219
QY	641	TCGAGAGCTCATGAAGGAGCAGGAGCTCTGCTGAGACCAAGCTCAGGTCTTACGAGAGGG	700
DB	1220	TGGAGAAGCAGCACCGGAGCAGAGCTGCAGCAGCTCAAGAACAGGAGAGGGCAAG	1279
QY	701	AGAGACCAAGCTTCGGCCCGCGCTGGAGAGA	733
DB	1280	AGAGTCCCTGGCCAGCACAGAAATGAAGATGA	1312
RESULT 14			
US-09-817-913-8			
; Sequence 8, Application US/09817913			
; Patent No. US20020061860A1			
; GENERAL INFORMATION:			
; APPLICANT: Li, Zuomei			
; APPLICANT: Bonfils, Claire			
; APPLICANT: Besterman, Jeffrey			
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms			
; FILE REFERENCE: 106101.145			
; CURRENT APPLICATION NUMBER: US/09/817,913			
; CURRENT FILING DATE: 2001-03-26			
; PRIOR APPLICATION NUMBER: US 60/192,157			
; PRIOR FILING DATE: 2000-03-24			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 8			

[illegible]

Search completed: June 15, 2003, 06:49:12
Job time : 119.039 secs

```

; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Human
US-09-817-913-8

Query Match          7.2%   Score 57.8;   DB 10;   Length 8459;
Best Local Similarity 48.3%;   Pred. No. 4.3e-05;
Matches 161;   Conservative 0;   Mismatches 172;   Indels 0;   Gaps 0.

Qy  401 AGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAAGTGCAGCGCAGCTTTGAGGAGAAAGGAGC 460
Db  980 AGCGCGCCCTGCGGGAGCAGCAGCTGTCAGCAGAGAGTCTCTGGCGCTCAAGCAGAAAGCAGC 1039
Qy  461 TTGCCTCCAGCCTTGGCCCTACGAGAGGGCGCGGGCGCTGCAGGGACGAGCTTGAGGGCC 520
Db  1040 AGATCCAGAGGCAGATCTCTATCGCCAGTTCAGAGGCGACGACGAGCTCTCCCGC 1099
Qy  521 CGAGGCCCAAGGGGGGCAACAGCTCAAGCAGGCTCGCGAGAGAGCCAGCGCGCGCAGC 580
Db  1100 AGCACGAGGGCGCAGCTCCACGAGCAGCATCAAGCAATTAACAGGAGATGCTCGCCATGAAGC 1159
Qy  581 AGGTCTCTGCACCTGCGAGGTACTGCAGCTTCAGCAGAGAGAACGGCGAGCTCCGGCAGGAGC 640
Db  1160 ACCAGCAGGAGCTGCTGTAACACACAGCGGAAGCTTGGAGAGGCCACCGCCAGGAGCAGGAGC 1219
Qy  641 TCGAGAGCCTCATGAAGGAGCAGGACCTGCTGGAGACCAGCTTCAGTCTCTACGAGAGGG 700
Db  1220 TGGAGAAGCAGCACCGGGAGCAGAGCTGCAGCAGCTCAAGAACAGAGGAGAGGGCAAG 1279
Qy  701 AGAAGACCAAGCTTGGCCCCCGCTGGAGAGA 733
Db  1280 AGAGTGCCTGGCCGAGCAGCAGAAAGTGAAGATGA 1312

```

```

RESULT 15
US-09-817-538-8
; Sequence 8, Application US/09817538
; Patent NO. US20020137162A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfills, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
; TITLE OF INVENTION: Deacetylase Isoforms
; FILE REFERENCE: 106101.144
; CURRENT APPLICATION NUMBER: US/09/817,538
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Human
US-09-817-538-8

```

	Query Match	7.2%	Score 57.8;	DB 10;	Length 8459;
	Best Local Similarity	48.3%;	Pred. No. 4.3e-05;		
	Matches 161;	Conservative 0;	Mismatches 172;	Indels 0;	Gaps 0;
Qy	401	AGCTGTTTGAGAGAGGAGGGCCCTCCAGAAAGCTGCAGCAGCTTTTGAGGAAGAGGAGC	460		
Db	980	AGCCGGGCCCTGCGGGAGCAGCAGCTGCAGCAGAGAGCTCTTGGCGCTCAAGCAGAAGCAGC	1039		
Qy	461	TTGCGCTCCAGCCTGGCCCTACGAGAGGCGCGCGCGCTGCAGGGACGAGCTGGAGGGCC	520		
Db	1040	AGATCCAGAGGCAGATCCTCATGCCCGATTCCAGAGGCACGACGAGCTCTCCCGGC	1099		
Qy	521	CGGAGCCCAAGGGCGGCAACAGCTCAAGCAGGCCTTCGAAAGAGCCAGCGCGCGCAGC	580		
Db	1100	AGCACGAGGGCGAGCTCCAGCAGCACATCAAGCAATAACAGGAGATGCTGCACATGAAGC	1159		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 1119.4 Seconds
(without alignments)
11632.279 Million cell updates/sec

Title: US-09-513-888c-1_COPY_1707_2510

Perfect score: 804

Sequence: 1 gccatctgactctctcccc.....tttggcgccagtaccacct 804

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	507.2	63.1	740	14 BQ769435	BQ769435 UI-M-FIO-
2	468.6	58.3	547	12 BF590813	BF590813 7h42e07.x
3	463.4	57.6	545	10 AW007737	AW007737 wt68e06.x
4	448.8	55.8	668	12 BF058214	BF058214 7K28d05.x
5	445.4	55.4	447	9 A1042490	A1042490 ox62e04.x
6	429.8	53.5	817	9 A1984777	A1984777 wr85b07.x

C	7	365	45.4	397	10	AW028197	AW028197
C	8	336.6	41.9	344	10	AW016544	AW016544 UI-H-BIOP
C	9	296.8	36.9	743	10	BE410921	BE410921 601303579
C	10	285.2	35.5	673	10	BE384131	BE384131 601272956
C	11	254.8	31.7	407	10	AV663617	AV663617 AV663617
C	12	246	30.6	394	12	BF549120	BF549120 UI-R-AO-a
C	13	245	30.5	394	12	BF554804	BF554804 UI-R-E0-c
C	14	179	22.3	725	9	AJ454524	AJ454524 AJ454524
C	15	172.4	21.4	531	10	BE276168	BE276168 601144424
C	16	151.4	18.8	153	9	AA644373	AA644373 af64c10.8
C	17	144.4	18.0	596	10	AV663618	AV663618 AV663618
C	18	135.8	16.9	926	12	BF583930	BF583930 602098941
C	19	125.4	15.6	907	14	BQ652760	BQ652760 AGENCOURT
C	20	125.4	15.6	921	14	BQ652556	BQ652556 AGENCOURT
C	21	123.8	15.4	991	14	BQ720762	BQ720762 AGENCOURT
C	22	123.4	15.3	941	14	BQ955592	BQ955592 AGENCOURT
C	23	121.2	15.1	635	-12	BF031137	BF031137 601456991
C	24	115	14.3	924	12	BE897897	BE897897 601440570
C	25	113.8	14.2	940	14	BQ895397	BQ895397 AGENCOURT
C	26	112.4	14.0	561	9	AA875041	AA875041 UI-R-E0-c
C	27	112.2	14.0	1101	14	BQ051027	BQ051027 AGENCOURT
C	28	109.8	13.7	460	12	BE757942	BE757942 212520 MA
C	29	106.2	13.2	647	12	BG338379	BG338379 602436164
C	30	105.4	13.1	700	13	B1854892	B1854892 603381873
C	31	104.6	13.0	663	9	AL635635	AL635635 AL635635
C	32	104.6	13.0	1115	12	BF784037	BF784037 602107882
C	33	104.4	13.0	735	12	BE740090	BE740090 601595016
C	34	104.2	13.0	906	13	BG917774	BG917774 602818743
C	35	100.6	12.5	487	13	B1338021	B1338021 361664 MA
C	36	99	12.3	1093	14	BM811378	BM811378 AGENCOURT
C	37	98.6	12.3	661	12	BF312957	BF312957 601896225
C	38	94	11.7	602	14	BM717087	BM717087 UI-E-EJO-
C	39	94	11.7	603	14	BM674468	BM674468 UI-E-EJO-
C	40	92.4	11.5	723	10	BE312971	BE312971 601150210
C	41	90.4	11.2	806	14	BM947278	BM947278 UI-M-EHOP
C	42	89.6	11.1	932	14	BQ879505	BQ879505 AGENCOURT
C	43	87.8	10.9	846	12	BF970582	BF970582 602274088
C	44	87.8	10.9	923	12	BG750395	BG750395 602709229
C	45	87.8	10.9	1298	13	BM470160	BM470160 AGENCOURT

ALIGNMENTS

RESULT 1
BQ769435
LOCUS BQ769435 740 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-M-FIO-byq-a-22-0-UI.r1 NIH_BMAP_FIO Mus musculus cDNA clone
IMAGE:6400797 5', mRNA sequence.
ACCESSION BQ769435
VERSION BQ769435.1 GI:21977909
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 740)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BNAP)
Seq primer: pYX-5.
Location/Qualifiers

FEATURES

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 817)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1307 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 428.

FEATURES

Location/Qualifiers

1..817

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2494453"

/clone_lib="NCI-CCAP_Kid11"

/lab_host="DH10B"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI-CCAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 132376-132391, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 133 a 235 c 266 g 180 t 3 others

Query Match 53.5%; Score 429.8; DB 9; Length 817;
Best Local Similarity 97.6%; Pred No. 4.7e-74;
Matches 457; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCACACACAGCTGCACCCGCGCTCCAGAC 60
Db GCCATCTGCACTCTCCCGGAGAGTGCACACACAGCTGCACCCGCGCTCCAGAC 409
Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTCTGGGGGCTCTCAGACTCCGGC 120
Db AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTCTGGGGGCTCTCAGACTCCGGC 408
Qy 121 CGGAATCTCATGTCCAGCTGCCACACACAG-CACACAGCAGCAGCTACCAGCT--GGAC 177
Db CGGAATCTCATGTCCAGCTGCCACACACAGCAGCAGCTACCAGCTACCAGCTGGGAC 289
Qy 178 CGCTGTGTACACCCGCTGGGACCCACAAAGCGTTTGGGGGCTCGGCCACAAATCACC 237
Db CGCTGTGTACACCCGCTGGGACCCACAAAGCGTTTGGGGGCTCGGCCACAAATCACC 229
Qy 238 CAGGCACTGTCTCCAGCAGCAGCAATGATGAGCTTGAAGCTCTCTCTTCTCCGAC 297
Db CAGGCACTGTCTCTCCAGCAGCAGCAATGATGAGCTTGAAGCTCTCTCTTCTCCGAC 228
Qy 298 GGAGGTAGCAAGCTGGGGCACTCGAAACAGGAGGAGGCGCTCTGTGTCTCGCTCC 357
Db GGAGGTAGCAAGCTGGGGCACTCGAAACAGGAGGAGGCGCTCTGTGTCTCGCTCC 109
Qy 358 CCCATCTCCACCGAGCTGAGCATCCAGAGCTGGAAACAGAGCTGTTGGAGGGAG 417
Db CCCATCTCCACCGAGCTGAGCATCCAGAGCTGGAAACAGAGCTGTTGGAGGGAG 49
Qy 418 GGCGCCCTCCAGAGCTCAGCGAGCTTTAGGAGAGGAGCTTGCC 465

Db 48 GCGCGCTCCAGAAAGCTGCAGCGAGCTTTGAGGAGAGGAGCTTGCC 1

RESULT 7
LOCUS
DEFINITION
Accession
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW028197 397 bp mRNA linear EST 09-MAR-2000
W84801.X1 Soares_thymus_NHfTh Homo sapiens cDNA clone
IMAGE:2536200 3' Similar to TR:060299 O60299 KIAA0552 PROTEIN. ;
mRNA sequence.
AW028197
AW028197.1 GI:5886953
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 397)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1067 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 139.

FEATURES
Location/Qualifiers
1..397
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2536200"
/clone_lib="Soares_thymus_NHfTh"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus, pooled; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTACCAATCTGAGTGGAGCGCGCGACAGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 67 a 120 c 124 g 86 t

Query Match 45.4%; Score 365; DB 10; Length 397;
Best Local Similarity 95.0%; Pred. No. 1.7e-61;
Matches 377; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 97 TCTGGGCGGTGTCTCAGACTCCGGCGGAACTCCATGTCAGCCTGCCACACAGACACC 156
Db TCTGGGCGGTGTCTCAGAACTCCGCTTGAATTCATGTCAGCCTGCCACACAGACACA 338
Qy 157 AGCAGCAGCTACCAAGCTGGAGCGGCTGTACACCCGCTGGAGCCCAAGCCGTTTGGG 216
Db GGCAGCAGCTACCAAGTGGAGCGGCTGTGTACACAGTGGAGCCCAAGCCGTTTGGG 278
Qy 217 GGCTCCGCCCCACAAATCACCAGGGCATCGTCTCCAGGACAGCAACATGATGAGCCTG 276
Db GGCTCCGCCCCACAAATCACCAGGGCATCGTCTCCAGGACAGCAACATGATGAGCCTG 218
Qy 277 AAGCTCTGTCTCTTCTCCGAGCGAGGTAGCAAGCTGGGCCCACTCGAAACAGGACAGCAAG 336
Db AAGCTCTGTCTCTTCTCCGAGCGAGGTAGCAAGTGGGCCCACTCGAAACAGGACAGCAAG 158
Qy 337 GGCCCTCTCGTGTGTCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 396
Db GGCCCTCTCGTGTGTCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 98

397 CAGAGCTGTTGGAGAGGAGGGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAAG 456
 |||||
 Db 97 CAGAGCTGTTGGAGAGGAGGGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAAG 38
 |||||

457 GAGCTTGCTCCAGCTGCGCTACGAGGAGCGGCGC 493
 |||||
 Db 37 GAGCTTGCTCCAGCTGCGCTACGAGGAGCGGCGC 1
 |||||

RESULT 8
 AW016544/c
 LOCUS
 DEFINITION
 UI-H-B10p-abg-g-04-01.s1 NCI_CGAP_Sub2 Homo sapiens cDNA clone
 IMAGE:2711982 3, mRNA sequence.
 ACCESSION
 AW016544
 VERSION
 AW016544.1 GI:5865301
 KEYWORDS
 EST.
 SOURCE
 HOMO sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 344)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: NCI-CGAP clone distribution
 information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA=No.

Location/Qualifiers
 1. 344
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2711982"
 /clone_lib="NCI CGAP Sub2"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTY3D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not 1; Site 2: Eco RI; The
 NCI_CGAP_Sub2 library is a subtracted library derived from
 B1. B1 constitutes a mixture of 21 normalized or
 subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP_Pr22,
 NCI CGAP_Pr28, NCI CGAP Co10, NCI CGAP Co16,
 NCI CGAP_Kid5, NCI CGAP Kid12, NCI CGAP Kid3,
 NCI CGAP_Kid11, NCI CGAP_Lym2, NCI CGAP_Br2, NCI CGAP_Co8,
 NCI CGAP_Cu11, NCI CGAP_Le12, NCI CGAP_Brn23, NCI CGAP_Lu5,
 NCI CGAP_Lu24, NCI CGAP_Lu19, NCI CGAP_GC4, NCI CGAP_GC6,
 NCI CGAP_Brn25. These 21 libraries were pooled and a
 single-stranded DNA preparation of the resulting mixture
 was used as a tracer in a subtractive hybridization with a
 driver whose composition is detailed below: NCI CGAP Kid3
 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE_
 Clonids 132376-132391), 1456008-1456775, 1500552-1502855
) NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725,
 3776-3778 (IMAGE Clonids 1323912-1325831,
 1471368-1472903, 1492104-1493255) NCI CGAP Lu5 pool 1 LLAM
 3575-3582, 3851-3854 (IMAGE Clonids 1414920-1417991,
 1520904-1522439) NCI CGAP GC4 pool 1 LLAM 3164-3167,
 3716-3720, 3733-3735 (IMAGE Clonids 1257096-1258631,
 1469064-1470983, 1475592-1476743) NCI CGAP Pr22 pool 1
 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonids
 985608-986759, 1101192-1101959, 1217928-1220615)
 NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
 Clonids 1057416-1061255, 1144584-1145351) The resulting
 subtracted library contained 4 million recombinants.
 Subtraction was performed as previously described (Bonaldi
 & Lennon & Soares (1996): Normalization and Subtraction:
 Two Approaches To Facilitate Gene Discovery. Genome

Research 6, 791-806.
 TAG LIB=NCI CGAP GC4
 TAG_TISSUE=germ Cell
 TAG_SBO=AAATC"

BASE COUNT 51 a 106 c 112 g 74 t 1 others
 ORIGIN

Query Match 41.9%; Score 336.6; DB 10; Length 344;
 Best Local Similarity 99.5%; Pred. No. 5.8e-56;
 Matches 339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 149 ACAGCACCAGCAGCAGCTACCGTGGACCGCTGGTGCACACCCGCTGGGACCCACAAGCC 208
 |||||
 Db 344 ACAGCACCAGCAGCAGTTACCGTGGACCGCTGGTGCACACCCGCTGGGACCCACAAGCC 285
 |||||

Oy 209 GTTTTGGGGGCTCGCCCAACAATCACCAGGCGATCGTCTCCAGGACGAGCAACATGA 268
 |||||
 Db 284 GTTTTGGGGGCTCGCCCAACAATCACCAGGCGATCGTCTCCAGGACGAGCAACATGA 225
 |||||

Oy 269 TGAGCCTGAAGGCTCTGCTCTTCCGACGAGGTAGCAAGCTGGGCGACCTCGAACAAGG 328
 |||||
 Db 224 TGAGCCTGAAGGCTCTGCTCTTCCGACGAGGTAGCAAGCTGGGCGACCTCGAACAAGG 165
 |||||

Oy 329 CAGACAAGGGCCCTCGTGTGTCGGCTCCGCCATCTCCAGGACGAGTGCAGCATCCAGG 388
 |||||
 Db 164 CAGACAAGGGCCCTCGTGTGTCGGCTCCGCCATCTCCAGGATGATGCGAGCATCCAGG 105
 |||||

Oy 389 AGCTGGAACAGAGCTGTTGGAGAGGAGGGCGGCCCTCCAGAGCTCGAGCGAGCTTTG 448
 |||||
 Db 104 AGCTGGAACAGAGCTGTTGGAGAGGAGGGCGGCCCTCCAGAGCTCGAGCGAGCTTTG 45
 |||||

Oy 449 AGGAGAAGGAGCTTGGCTCCAGCGCTCGCTTACGAGGAGCGCGC 492
 |||||
 Db 44 AGGAGAAGGAGCTTGGCTCCAGCGCTCGCTTACGAGGAGCGCGC 1
 |||||

RESULT 9
 BE410921/c
 LOCUS
 DEFINITION
 601303579F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638203 5',
 mRNA sequence.
 ACCESSION
 BE410921
 VERSION
 BE410921.1 GI:9347371
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 743)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM338 row: m column: 20
 High quality sequence start: 49
 High quality sequence stop: 688.
 Location/Qualifiers
 1. 743
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3638203"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_hosts="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pORF7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.

	119 a	229 c	226 q	168 t	1 others
(categories) and superscripts in KT (file technologies).					

	119 a	229 c	226 q	168 t	1 others
(categories) and superscripts in KT (file technologies).					

36.9%;	Score	296.8;	DB 10;	Length	743;
Similarity	97.0%;	Pred. No.	3.9e-48;		
Conservative	0;	Mismatches	8;	Indels	2; Gaps 2;
SCCTCGGCCCAACAATCACCCAGGCGATCGTCTCCAGGACAGCAACATGATGAGCGCTG	216				
CGGCGCGGCCCGAAGAACACAGG-CATCGTCTCTCCAGGACGCAACATGATGAGCGCTG	653				
AGGCTCTGTCTTCTCCGAGCGAGGTAGCAAGCTGGGCCACTCGAAACAAGGCGACACAAG	336				
A-GCTCTGTCTTCTCCGAGCGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACACAAG	594				
CCCCCGCTGTGTGTCCGCTCCCCCATCTCCACGGACGAGTGCAGACTCCAGGAGCTCGAA	396				
CCCCCGCTGTGTGTGTCCGCTCCCCCATCGCCACGGACGAGTGCAGACTCCAGGAGCTGGAG	534				
CGAGAGCTGTGGAGAGGAGGGCGCCCTCCAGAGACTGCAGCCGACGCTTTGAGGAGAAG	456				
AGAAAGCTGTGGAGAGGAGGGCGCCCTCCAGAAAGTTCAGACCGCAGCTTTGAGGAGAAG	474				
AGCTTTGCTTCAGGCTTGGCTCTACGAGGAGCGGCGCGGCGCTCGAGGACGAGCTCGAG	516				
AGCTTTGCTTCAGGCTTGGCTCTACGAGGAGCGGCGCGGCGCTCGAGGACGAGCTCGAG	414				
CCCCCGGAGCCCAAGGCGGCAACAAGCTCAAG	549				
CCCCCGGAGCCCAAGGCGGCAACAAGCTCAAG	381				

044131 673 bp mRNA linear EST 21-JUL-2000
7272956F1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:3614076 5',
a sequence.
044131
044131.1 GI:9329496

sapiens
 Vertebrata; Chordata; Metazoa; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 673)
 MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 published (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Reagent Procurement: ATCC/DC/PT/DTF
 Library Preparation: Ling Hong/Rubin Laboratory
 Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Sequencing by: Incyte Genomics, Inc.
 Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Locus: LLCM275 row: p column: 13
 quality sequence stop: 670.

```

1 quality sequenceObj: 970.
location/Qualifiers
    1..673
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3614076"
    /clone_lib="NIH_MGC_20"
    /tissue_type="melanotic melanoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: skin; Vector: pOTB3; Site 1: hOx1; Site 2:

```

EcORI, cDNA made by oligo-dT priming. Directionally
 cloned into pSRI/XhoI sites using the following 5',
 adaptor GGCACG(C)G. Sizes selected >500bp for average
 insert size 1 kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 110 a 211 c 208 g 144 t
 BASE COUNT

Query Match	35.5%;	Score 285.2;	DB 10;	Length 673;	
Best Local Similarity	98.4%;	Pred. No. 7.1e-46;			
Matches 309;	Conservative 0;	Mismatches 3;	Indels 2;	Gaps 2;	
Qy	237	CCAGGCGATCGTCTCCAGGACAGCAACATGATGAGCGCTGAAGGCTGTGCTTCTCCGA	296		
Db	669	CCAGGCGATCGTCTCCA-GACAGCAACATGATGAGCGCTGTGCTTCTCCGA	611		
Qy	297	CGGAGGTAGCAAGCTGGGCCACTCGAAACAAGGCAGACAAGGG-CCGCTCGTGTGTCCGCT	355		
Db	610	CGGAGGTAGCAAGCTGGGCCACTCGAAACAAGGCAGACAAGGGCCCCCTCGTGTGTCCGCT	551		
Qy	356	CCCCCATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAACAGAGAGCTTTGGAGAGGG	415		
Db	550	CCCCCATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAAGCTTTGGAGAGGG	491		
Qy	416	AGGGCGCCCTCCAGAGAGCTGCAGCGCAGCTTTTTCAGGAGAGGAGCTTGCTCCAGCCTGG	475		
Db	490	AGGGCGCCCTCCAGAGAGCTGCAGCGCAGCTTTTTCAGGAGAGGAGCTTGCTCCAGCCTGG	431		
Qy	476	CTTACGAGGAGCGGCGCGCGCTGTCAGGGA	CGAGCTGGAGGGCCCGGAGCCCAAGGCG	535	
Db	430	CCTACGAGGAGCGGCGCGCGCTGTCAGGGA	CGAGCTGGAGGGCCCGGAGCCCAAGGCG	371	
Qy	536	GCAACAAGTCTAAG	549		
Db	370	GCAACAAGTCTAAG	357		

RESULT 11	
AV663617/c	
LOCUS	407 bp mRNA linear EST 28-NOV-2001
DEFINITION	AV663617 Bos taurus brain fetus Bos taurus cDNA clone E1BR037707
ACCESSION	AV663617
VERSION	3', mRNA sequence.
KEYWORDS	AV663617.1 GI:9922647
SOURCE	EST.
ORGANISM	COW.
	Bos taurus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Euteria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
	Bovidae; Bovinae; Bos.
REFERENCE	1 (bases 1 to 407)
AUTHORS	Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H. and Sugimoto,Y.
TITLE	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
JOURNAL	Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE	21570554
COMMENT	Contact: Yoshikazu Sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics Odakura, Nishigo, Niishi-shirakawa, Fukushima 961-8061, Japan Tel: 81-248-25-5641 Fax: 81-248-25-5725 Email: kazusugi@cococ.ocn.ne.jp Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library.
FEATURES	Location/Qualifiers
source	1..407 /organism="Bos taurus" /db_xref="tracv.9912"


```

/clone="E1BR037F07"
/clone_lib="Bos taurus brain fetus"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT      58 a   119 c   141 g   89 t
ORIGIN

Query Match      31.7%; Score 254.8; DB 10; Length 407;
Best Local Similarity 84.8%; Pred. No. 5.4e-40;
Matches 324; Conservative 0; Mismatches 47; Indels 11; Gaps 3;

QY 64 CCCAAGGAGCAGGAGCTGAAGCCTGGCTGTGCTCTGGGGCG--CTGTCAAGCTCCGGCC 121
Db 373 CCAAAGGAGCAGGAGCTGAAGCCTGGCTGTGCTCTGGGGCGCACTGTCCGAACCTCCGGCC 314
QY 122 GGAATCTCATGTCTCCAGCTGCCACACACAGCAGCAGCAGCTACCAGCTGGACCGC 181
Db 313 GGAATCTCATGTCTCCAGCTGCCACACAGCAGCAGCAGCTACCAGCTGGACCGC 254
QY 182 TGGTCACACCCCTGGAGCCCAAGCCGTTTGGGGGCTCCGCCACACATCACCCAGG 241
Db 253 TGGTCACCTCTGTGGGGCCGCGCCGCTTGGGGGCTCAGCCACACATCACACAGG 194
QY 242 GCATCGTCTCCAGACAGCAACATGATGAGCTTAAGGCTGTGCTTCTCCGACGAG 301
Db 193 GCATGTGTCTCCAGACAGCAACATGATGAGCTTAAGGCGCTGTCTTCTCTGACGGG 134
QY 302 GTAGCAAGCTGGCCACCTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
Db 133 GAGCAAGCTGGCCACCGCAAGCAAA---AGACAAGGGC-----TCTGTGGCTCCCGCA 83
QY 362 TCTTCAACGAGCAGTGCGAGCTCCAGGAGCTGGAACAGAGCTGTGGAGAGGAGGAGG 421
Db 82 TCTTCAACGAGCAGTGCGAGCTCCAGGAGCTGGAACAGAGCTGTGGAGAGGAGGAGG 23
QY 422 CCTTCCAGAGCTGACGGCGAG 443
Db 22 AGCTGCAGAGGCTGCACCGCAG 1

RESULT 12
BF549120
LOCUS
DEFINITION
UI-R-A0-af-c-04-0-UI.r1 UI-R-A0 Rattus norvegicus cdna clone
UI-R-A0-af-c-04-0-UI 5', mRNA sequence.
ACCESSION
BF549120
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 394)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
9704477
MEDLINE
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1791781

Seq primer: M13 Forward.
FEATURES
Location/Qualifiers
1..394
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="UI-R-A0-af-c-04-0-UI"
/clone_lib="UI-R-A0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; This library
consists of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung, brain
liver, kidney, heart, spleen, ovary, and muscle. The tag
is a string of 3-5 nucleotides present between the Not I
site and the oligo-dT track which allows identification of
the library of origin of a clone within the mixture."
BASE COUNT      89 a   125 c   103 g   77 t
ORIGIN

Query Match      30.6%; Score 246; DB 12; Length 394;
Best Local Similarity 85.8%; Pred. No. 2.8e-38;
Matches 273; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCAGCTGCACCCGCCCTCCAGAC 60
Db 76 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCAGCTGCACCCGCCCTCCAGAC 135
QY 61 AAGCCCAAGCAGGAGGAGCTGAAGCCTGGGCTGTGCTGTGGGGCGCTGCAGACTCCGGC 120
Db 136 AAGCCCAAGCAGGAGGAGCTGAAGCCTGGGCTGTGCTGTGGGGCGCTGCAGACTCCGGC 195
QY 121 CGGAATCTCATGTCTCCAGCTGCCACACAGCAGCAGCAGCTACAGCTGCAGCCCG 180
Db 196 CGGAATCTCATGTCTCCAGCTGCCACACAGCAGCAGCAGCTACAGCTGCAGCTCCT 255
QY 181 CTGCTCACACCCGCGGAGCCCAAGCGGTTTGGGGGCTCCGCCACACATCACCCAG 240
Db 256 CTGCTCACACCCGCGGAGCCCAAGCGGTTTGGGGGCTCCGCCACACATCACCAA 315
QY 241 GGATCTCTCTCCAGGACAGCAACATGATGAGCTGAAGGCTCTGTCTTCTCCGAGGA 300
Db 316 GGATCTCTCTCCAGGACAGTAATATGATGAGCTGAAGGCTCTGTCTTCTCTGATGG 375
QY 301 GGTAGCAAGCTGGGCCAC 318
Db 376 GGCAGCAAGCTGGGCTCAC 393

RESULT 13
BF554804
LOCUS
DEFINITION
UI-R-E0-cb-b-03-0-UI.r1 UI-R-E0 Rattus norvegicus cdna clone
UI-R-E0-cb-b-03-0-UI 5', mRNA sequence.
ACCESSION
BF554804
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 394)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
9704477
MEDLINE
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1791781

```


Tel: 319 335 8250
Fax: 319 335 9565

Email: msoaree@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1770031
Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
1..394

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-E0-cb-b-03-0-UI"
/clone_lib="UI-R-E0"
/dev_stage="embryonic"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTY3D-Pac (Pharmacia) with a modified
polylinker; Site_1: NotI; Site_2: EcoRI; This library
consists of a mixture of individually tagged normalized
libraries constructed from 8, 12 and 18-day embryo. The
tag is a string of 3-5 nucleotides present between the
Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture."

BASE COUNT 89 a 125 c 102 g 77 t 1-others

Query Match 30.5%; Score 245; DB 12; Length 394;

Best Local Similarity 85.5%; Pred. No. 4.3e-38;
Matches 272; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 GCCATCTCGACTCTCCCGGAGAGTGCACGACACACAGCTGACCCCGCCCTCCAGAC 60
DB 76 GCCATCTCGACTCTTCACTGAGAGTACCAACACAGCTTCACTCCATGCTCCAGAT 135
QY 61 AAGCCCAAGGAGCAGAGCTGAGCTGGCTGTGCTGTGGGGCTGTGACACTCCGGC 120
DB 136 AAGCCCAAGGAGCAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 195
QY 121 CGGAATCCATGCTCCAGCTGCGCCACACACAGCAGCAGCAGCTACCGCTGAGACCG 180
DB 196 CGGAATCCATGCTGAGCTGCGCCACGATAGCACCAGCAGCTACCGCTGAGCTGAGCT 255
QY 181 CTGGTCACACCGCTGGGACCCACAGCCGTTTGGGGGCTCGGCCACCAATCAACCCAG 240
DB 256 CTGGTCACACCGCTGGGACCCACAGCCGTTTGGGGGCTCGGCCACCAATCAACCA 315
QY 241 GGCATCGCTCCAGGACGACGACATGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 300
DB 316 GGCATCGCTCCAGGACGACGACATGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 375
QY 301 GGTAGCAAGCTGGGCCAC 318
DB 376 GGCACCAAGCTGGCTCAC 393

RESULT 14

LOCUS AJ454524 725 bp mRNA linear EST 22-APR-2002
DEFINITION AJ454524 riken1 Gallus gallus cdna clone 3a8r1, mRNA sequence.

ACCESSION AJ454524

VERSION AJ454524.1 GI:20264620

KEYWORDS EST.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 725)

AUTHORS Buerstedde, J.M.

TITLE Gallus gallus bursal lymphocyte EST

JOURNAL Unpublished (2002)

COMMENT

Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..725

FEATURES

Location/Qualifiers
1..725

/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="3a8r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"

BASE COUNT 138 a 190 c 293 g 100 t 4 others

Query Match 22.3%; Score 179; DB 9; Length 725;

Best Local Similarity 66.8%; Pred. No. 3.7e-25;

Matches 255; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 383 TCCAGGAGCTGGAACAGAGCTGTTGGAGAGGGGCGCCCTCCAGAGCTGCAGCGCA 442
DB 68 TCGTGAGCTGGAACAGAGCTGCGGAGAGGGGCGGAGCTGCAGAGCTGCAGCGCA 127
QY 443 GCTTTGAGGAGAAGAGCTTGCCTCCAGCTGCGCTACGAGAGCGGCGCGCGCTGCA 502
DB 128 GCCTGGATGAGATGAGGTGGCCATCTGCCAGGTGTACGAGAGAGAGAGCGCTGCTGTG 187
QY 503 GGGAGAGCTGGAGGGCCCGAGGCCAAAGGGGGCAACAAGCTCAAGCAGGCTCGCAGA 562
DB 188 AGCAGGAGCTGGAGGGTTTGGCGCAGCGCTGTGCGGCCAGGTGCGCAGGCGGCCAGC 247
QY 563 AGAGCCAGCGCGCAGCAGCTGCTGACCTGCGAGTACTGAGCTTCCAGCTTCCAGCAGGAGAGC 622
DB 248 AGGCGCAGCTGGGAGGAGCTGCTGCGAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 307
QY 623 GGCAGCTCCGCGCAGAGCTCGAGAGCTTCATGAAGGAGCAGACCTGCTGGAGAGCAAGC 682
DB 308 AGCAGCTGAGGAGGAGCTTGTCTCAGCTGCTGAGAGCGGAGCTGCTGGAGCGCCGCT 367
QY 683 TCAGGTCTCAGAGAGGAGAGAGCAGCTTCCGCGCCCGCTGGAGAGAGCCAGTGGG 742
DB 368 GCGCATCTCTCCAGCGGAAACACAGGAGCTGGGGCCACGCTGGAGGAGAGCAAGTGGG 427
QY 743 AGGTGAGGCCACAGAGGCTCA 764
DB 428 AGGTGCGCAGAGTCCGGTGA 449

RESULT 15

LOCUS BE276168/c

DEFINITION 60114424P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3050716 5',
mRNA sequence.

ACCESSION BE276168

VERSION BE276168.1 GI:9151131

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 531)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LCM96 row: 0 column: 05
 High quality sequence stop: 531.
 Location/Qualifiers
 1. 531
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3050716"
 /clone_lib="NIH MGC 20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTS7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 87 a 173 c 156 g 115 t
 ORIGIN

Query Match 21.4%; Score 172.4; DB 10; Length 531;
 Best Local Similarity 99.4%; Pred. No. 6.6e-24;
 Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 376 TGCAGCATCCAGGAGCTGGACAGAAAGCTTTGGAGAGGAGGGCGCCCTCCAGAAAGCTG 435
 Db |||||
 QY 531 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTTTGGAGAGGAGGGCGCCCTCCAGAAAGCTG 472
 Db |||||
 QY 436 CAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCTACGAGGAGCGCGCGG 495
 Db |||||
 QY 471 CAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCTACGAGGAGCGCGCGG 412
 Db |||||
 QY 496 CGCTGAGGAGCAGCTGGAGGGCGCGGAGCCCAAGGCGGCAACAAGCTCAAG 549
 Db |||||
 QY 411 CGCTGAGGAGCAGCTGGAGGGCGCGGAGCCCAAGGCGGCAACAAGCTCAAG 358
 Db |||||

Search completed: June 15, 2003, 03:41:52
 Job time : 1123.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 1662.94 Seconds
(without alignments)
11183.052 Million cell updates/sec

Title: US-09-513-888C-1_COPY_4912_5550
Perfect score: 639
Sequence: 1 gaggtgaacgcgaaggtag.....tggaggcagggtcttcocg 639

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rtd.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	639	100.0	5492	9 AF123659	Homo sapi
2	639	100.0	9108	9 AF123653	Homo sapi
3	626.4	98.0	173264	2 AC025853	Homo sapi
4	445	69.6	227884	2 AC114995	Mus muscu
C 5	445	69.6	263546	2 AC099416	Mus muscu
C 6	438.6	68.6	191210	2 AC108987	Rattus no
7	434	67.9	1722	9 AF123658	Homo sapi
8	374	58.5	1692	9 AF123657	Homo sapi
C 9	232.4	36.4	85530	2 AC125907	Rattus no
10	224.4	35.1	1614	9 AF123655	Homo sapi
11	175.4	27.4	1515	9 AF123656	Homo sapi
12	148.2	23.2	1612	9 BC005855	Homo sapi
13	148.2	23.2	2275	9 HSM805394	Homo sapi
14	148.2	23.2	2766	9 BC006212	Homo sapi
15	148.2	23.2	5733	9 AB058716	Homo sapi
16	148.2	23.2	181086	9 AL133215	Human DNA
17	143.6	22.5	1935	9 AY029201	Homo sapi
18	140.4	22.0	2099	9 AB046013	Macaca fa
19	136.2	21.3	2767	10 BC014695	Mus muscu
20	125.2	19.6	5257	9 AB011124	Homo sapi
C 21	125.2	19.6	125856	9 HS1187M17	Human DNA
C 22	123.4	19.3	103729	2 AC096620	Continuation (4 of
23	123.4	19.3	270700	2 AC105485	Rattus no
C 24	120.2	18.8	168210	2 AC116700	Mus muscu
C 25	118.2	18.5	167636	2 AC107097	Rattus no
C 26	118.2	18.5	225045	2 AL833803	Mus muscu
C 27	116.8	18.3	184865	2 AL807824	Mus muscu
C 28	116.8	18.3	207683	2 AC098712	Mus muscu
C 29	116	18.2	195673	2 AL807807	Mus muscu
C 30	115.2	18.0	287323	2 AC093363	Mus muscu
31	114.6	17.9	218760	2 AC105989	Mus muscu
32	114.4	17.9	1150	14 HS4ULIR3	epstein-bar
33	114.4	17.9	1926	6 AX107940	Sequence
34	114.4	17.9	2580	6 AR108994	Sequence
C 35	114.4	17.9	5452	6 AR083151	Sequence
36	114.4	17.9	5452	12 U02454	Cloning vec
37	114.4	17.9	9600	6 A92665	Sequence
38	114.4	17.9	9600	6 AR158345	Sequence
39	114.4	17.9	10596	6 I25041	Sequence
40	114.4	17.9	10596	6 I30503	Sequence
41	114.4	17.9	10737	12 XXU02428	Cloning vec
42	114.4	17.9	10850	12 U02455	Cloning vec
43	114.4	17.9	172281	14 EBV	Epstein-Bar
44	114.4	17.9	184113	14 HS4B958RAJ	Epstein-Bar
C 45	114.2	17.9	155662	2 AC112388	Rattus no

ALIGNMENTS

RESULT 1
AF123659
LOCUS AF123659 5492 bp mRNA linear PRI 07-APR-1999
DEFINITION Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.
ACCESSION AF123659
VERSION AF123659.1 GI:4572475
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 5492)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

JOURNAL
MEDLINE
PUBMED
99199287
10097140

REFERENCE
AUTHORS
Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

TITLE
Direct Submission

JOURNAL
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA

FEATURES
Location/Qualifiers
1..5492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
1..5492
/map="8p22"
/gene="FEZ1"
112..1902
/gene="FEZ1"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23840.1"
/db_xref="GI:4572476"
/translation="MGSVSSLLISGHSHKCRASQYKLRKSHLLKLNRYSDGLLRF
PFLMPFNSQLEMGSEKAVPTAFKPVLPKRGSHHPDYTALSSDGLGQAGVDFPST
ELKPLGSGALSDGRNSMSLTHSTSSYQLDPLVTPVGTFRFGSAHNTQGI
LQDNMMSLKALSFSDGSKLHGNKADKPSVRSPISTDECSTQLEQLKLLREGA
LQKLSRFEELKASSLAYERPRRCDELEGPCKGNKLOKASOKSQAQOVLHLQ
VLQOQEKRLQRLQESLMKQDLETLKLSRYEREKTSFGPALEETOWEVCQKSEIS
LNKQKQSETEVNAKASEILGLKALQKDRKLEGLRLTQDLEGALRTKGLELVC
ENELQKNEALLREKVNLLQELQELRAQALARDMGPTTFPDEVPALELEBLR
AELREERQHDQMSGFQHERLVWKEEKVITYQKQLQSYVAMYQNRQLEKALQ
LARGSAGEPLEVDLEGADIPYEDIIATEI"

BASE COUNT 1137 a 1704 c 1565 g 1086 t

ORIGIN

Query Match 100.0%; Score 639; DB 9; Length 5492;
Best Local Similarity 100.0%; Pred. No. 7e-78;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGTGAACGCAAGCTAGCAGATCTCGGTCTCAAGGCACACTGAAGACACGGG 60
Db 1321 GAGGTGAACGCAAGCTAGCAGATCTCGGTCTCAAGGCACACTGAAGACACGGG 1380

QY 61 GCGAAGCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 120
Db 1381 GCGAAGCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 1440

QY 121 GGCCTGGAGCTGGAGTCTGTGAGAAATGAGCTGCGAGCGCAAGAACGAGCGGAGCTG 180
Db 1441 GGCCTGGAGCTGGAGTCTGTGAGAAATGAGCTGCGAGCGCAAGAACGAGCGGAGCTG 1500

QY 181 CTGGCGGAGAAAGTGAACCTGTGGAGCAGGAGCTGCAGAGCTGGGGCCCGAGCCGCC 240
Db 1501 CTGGCGGAGAAAGTGAACCTGTGGAGCAGGAGCTGCAGAGCTGGGGCCCGAGCCGCC 1560

QY 241 CTGGCCCGCGACATGGGGCGCGCCACCTTCCCGAGGAGCTGCCCTGCCCTGCGAGCGGAG 300
Db 1561 CTGGCCCGCGACATGGGGCGCGCCACCTTCCCGAGGAGCTGCCCTGCCCTGCGAGCGGAG 1620

QY 301 CTGGAGCGGCTCGGGCCCGAGCTCGGGAGGAGCGCGCAAGGCCATGACAGATGTCCTCG 360
Db 1621 CTGGAGCGGCTCGGGCCCGAGCTCGGGAGGAGCGCGCAAGGCCATGACAGATGTCCTCG 1680

QY 361 GCCTTCCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAAGAGTGTATTCAGTACCAG 420
Db 1681 GCCTTCCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAAGAGTGTATTCAGTACCAG 1740

QY 421 AAACAGCTGACGAGAGCTACGTGGCCCATGTACACAGCGGAACAGCGCTGGAGAGGCC 480
Db 1741 AAACAGCTGACGAGAGCTACGTGGCCCATGTACACAGCGGAACAGCGCTGGAGAGGCC 1800

QY 481 CTGCAGCAGCTGGCAGCTGGGGCAGCGCCGGGAGCCCTTGGAGTTGACCTGGAAGGG 540
Db 1801 CTGCAGCAGCTGGCAGCTGGGGCAGCGCCGGGAGCCCTTGGAGTTGACCTGGAAGGG 1860

QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGTGCTCTGGGAAGGC 600
Db 1861 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGTGCTCTGGGAAGGC 1920

QY 601 GAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 639
Db 1921 GAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 1959

RESULT 2
AF123653 9108 bp DNA linear PRI 07-APR-1999
LOCUS Homo sapiens FEZ1 (FEZ1) gene, complete cds.
DEFINITION
ACCESSION AF123653
VERSION AF123653.1 GI:4572463
KEYWORDS
SOURCE Homo sapiens.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 9108)

REFERENCE
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

JOURNAL
MEDLINE
PUBMED
99199287
10097140
2 (bases 1 to 9108)
Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA
JOURNAL
FEATURES Location/Qualifiers
1..9108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/gene="FEZ1"
join(112..456,1707..2510,4912..5553)
/gene="FEZ1"
/product="FEZ1"
join(112..456,1707..2510,4912..5553)
/gene="FEZ1"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23834.1"
/db_xref="GI:4572464"
/translation="MGSVSSLLISGHSHKCRASQYKLRKSHLLKLNRYSDGLLRF
PFLMPFNSQLEMGSEKAVPTAFKPVLPKRGSHHPDYTALSSDGLGQAGVDFPST
ELKPLGSGALSDGRNSMSLTHSTSSYQLDPLVTPVGTFRFGSAHNTQGI
LQDNMMSLKALSFSDGSKLHGNKADKPSVRSPISTDECSTQLEQLKLLREGA
LQKLSRFEELKASSLAYERPRRCDELEGPCKGNKLOKASOKSQAQOVLHLQ
VLQOQEKRLQRLQESLMKQDLETLKLSRYEREKTSFGPALEETOWEVCQKSEIS
LNKQKQSETEVNAKASEILGLKALQKDRKLEGLRLTQDLEGALRTKGLELVC
ENELQKNEALLREKVNLLQELQELRAQALARDMGPTTFPDEVPALELEBLR
AELREERQHDQMSGFQHERLVWKEEKVITYQKQLQSYVAMYQNRQLEKALQ
LARGSAGEPLEVDLEGADIPYEDIIATEI"

BASE COUNT 2021 a 2628 c 2504 g 1952 t 3 others
ORIGIN

Query Match 100.0%; Score 639; DB 9; Length 9108;
Best Local Similarity 100.0%; Pred. No. 6.3e-78;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAGGTGAACGCGCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACGCTGAAGACACGCGG 60
Db 4972 GAGGTGAACGCGCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACGCTGAAGACACGCGG 5031
QY 61 GCGAAGCTGGAGGGCTGGAGTGAAGACCCAGGACCTGGAGGGCGCCCTGGCGACCAAG 120
Db 5032 GCGAAGCTGGAGGGCTGGAGTGAAGACCCAGGACCTGGAGGGCGCCCTGGCGACCAAG 5091
QY 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCGCAAGAAACGAGCGGAGCTG 180
Db 5092 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCGCAAGAAACGAGCGGAGCTG 5151
QY 181 CTGCGGGAAGAGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTCGCGGCCCGACGCCGCC 240
Db 5152 CTGCGGGAAGAGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTCGCGGCCCGACGCCGCC 5211
QY 241 CTGCGGCCCGCATGCGGCGCCGACCTTCCCGAGGAGCGTCCCTGCGCTGCGAGCGGAG 300
Db 5212 CTGCGGCCCGCATGCGGCGCCGACCTTCCCGAGGAGCGTCCCTGCGCTGCGAGCGGAG 5271
QY 301 CTGGAGCGGCTGCGGCGCGGAGCTGCGGAGGAGCGGCAAGGCGCATGACAGATGTCCTCG 360
Db 5272 CTGGAGCGGCTGCGGCGCGGAGCTGCGGAGGAGCGGCAAGGCGCATGACAGATGTCCTCG 5331
QY 361 GCGTTCAGCATAGCGGCTCGTGTGGAAGGAGGAGAGGAGAGGAGGATGATTCAGTACCAG 420
Db 5332 GCGTTCAGCATAGCGGCTCGTGTGGAAGGAGGAGAGGAGAGGAGGATGATTCAGTACCAG 5391
QY 421 AAACAGCTGCACAGCAGTACGTGCGCATGTACGAGCGGCAAGCAGCGCTGGAGAAGGCC 480
Db 5392 AAACAGCTGCACAGCAGTACGTGCGCATGTACGAGCGGCAAGCAGCGCTGGAGAAGGCC 5451
QY 481 CTGCAGCAGCTGGCAGCTGGGACAGCGCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 540
Db 5452 CTGCAGCAGCTGGCAGCTGGGACAGCGCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 5511
QY 541 GCTGCATCCCTACGAGGACATCATAGCCTAGATCTGAGGGGCTGCTTGGGAAGGC 600
Db 5512 GCTGCATCCCTACGAGGACATCATAGCCTAGATCTGAGGGGCTGCTTGGGAAGGC 5571
QY 601 GAGTCTGGGACCTGGCAGTGGAGGAGCGAGGCTCTCCCG 639
Db 5572 GAGTCTGGGACCTGGCAGTGGAGGAGCGAGGCTCTCCCG 5610

```

RESULT 3

```

AC025853
LOCUS AC025853 173264 bp DNA linear HTG 17-JUN-2002
DEFINITION Homo sapiens chromosome 8 clone RP11-353K12 map 8, *** SEQUENCING
IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC025853
VERSION AC025853.13 GI:21431202
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEPIN.
SOURCE human.
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 173264)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Collins, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Dodge, S., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Headford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J.,

```

```

Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Piban, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tasfayev, H., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

```

Direct Submission

```

Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 173264)

```

REFERENCE

AUTHORS

```

Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,
Cooke, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicoll, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Ratta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Teefaye, S.,
Theodore, J., Toham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

```

Direct Submission

```

Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 17, 2002 this sequence version replaced gi:21321864.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

COMMENT

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WBIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7454
Center clone name: 353_K_12

```

```

-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 39461: contig of 39461 bp in length
* 39462 39561: gap of 100 bp
* 39562 108347: contig of 68786 bp in length
* 108348 108447: gap of 100 bp
* 108448 173264: contig of 64817 bp in length.

```

FEATURES

source

```

1. 173264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-353K12"

```

```

/clone.lib="RPC1-11 Human Male BAC"
BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others
ORIGIN
Query Match 98.0%; Score 626.4; DB 2; Length 173264;
Best Local Similarity 99.7%; Pred. No. 1.7e-76;
Matches 638; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GAGGTGAACGCCAAGCTAGCGAGATCTGGGTCTCAAGGCACACTGAAGCACGCGG 60
Db 27212 GAGGTGAACGCCAAGCTAGCGAGATCTGGGTCTCAAGGCACACTGAAGCACGCGG 27271
QY 61 GCACAGCTGGAGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 120
Db 27272 GCACAGCTGGAGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 27331
QY 121 GGCCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAACGAGCGGAGCTG 180
Db 27332 GGCCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAACGAGCGGAGCTG 27391
QY 181 CTGCGGAGAGGTGAACCTGCTGAGCGAGGAGCTGAGGAGCTCGGGCCCGAGCCGCC 240
Db 27392 CTGCGGAGAGGTGAACCTGCTGAGCGAGGAGCTGAGGAGCTCGGGCCCGAGCCGCC 27451
QY 241 CTGCGCGCGACATGGGCGCGCCACCTTCCCGAGGAGCTGCCCTGCGAGCGGAG 300
Db 27452 CTGCGCGCGACATGGGCGCGCCACCTTCCCGAGGAGCTGCCCTGCGAGCGGAG 27511
QY 301 CTGAGCGCGCTCGCGGCGCGAGCTCGGGAGGAGCGGCAAGCCATGACAGATGTCCTCG 360
Db 27512 CTGAGCGCGCTCGCGGCGCGAGCTCGGGAGGAGCGGCAAGCCATGACAGATGTCCTCG 27571
QY 361 GCGTCCAGCATGAGCGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 27572 GCGTCCAGCATGAGCGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 27631
QY 421 AAACAGCTGCACAGAGCTACGTGGCCATGTACACGCGGAAACGAGCGCTGGAGAAGCC 480
Db 27632 AAACAGCTGCACAGAGCTACGTGGCCATGTACACGCGGAAACGAGCGCTGGAGAAGCC 27691
QY 481 CTGAGCGAGCTGGCGAGCTGGGACAGCGCGGGAGCGCCCTGG-AGGTGACCTGGAAGG 539
Db 27692 CTGAGCGAGCTGGCGAGCTGGGACAGCGCGGGAGCGCCCTGGNAGGTGACCTGGAAGG 27751
QY 540 GCGTACATCCCTAGGAGGACATCATAGCCACTGAGTCTGAGGGGCTGCTGGAGG 599
Db 27752 GCGTACATCCCTAGGAGGACATCATAGCCACTGAGTCTGAGGGGCTGCTGGAGG 27811
QY 600 CGAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 639
Db 27812 CGAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 27851

RESULT 4
AC114995 227884 bp DNA linear HTG 26-JUN-2002
LOCUS Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered
DEFINITION pieces.
ACCESSION AC114995
VERSION AC114995.3 GI:21592059
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 277884)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-171P5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 277884)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gordon,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,A., Kells,C., LaRocque,K., Lamazares,R., Jones,C.,
Kamat,A., Karatas,A., Kell,C., Lagocque,K., Lamazares,R., Jones,C.,
Lander,E., Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
MacLean,C., MacDonald,P., Major,J., Marquis,N.,
McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,D.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Rieback,M., Riley,R., Rise,C.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 227884)
REFERENCE
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gordon,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lagocque,K.,
Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Roetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 26, 2002 this sequence version replaced gi:21535965.
All repeats were identified using RepeatMasker.
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23473
Center clone name: 171 P 5
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 225162 bases at least Q40
Consensus quality: 226535 bases at least Q30
Consensus quality: 227117 bases at least Q20
Insert size: 225000; agarose-fp
Insert size: 227484; sum-of-contigs
Quality coverage: 8.6 in Q20 bases; agarose-fp
Quality coverage: 8.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 301: contig of 301 bp in length
 * 302 401: gap of 100 bp
 * 402 5831: contig of 5430 bp in length
 * 5832 5931: gap of 100 bp
 * 5932 173029: contig of 167098 bp in length
 * 173030 173129: gap of 100 bp
 * 173130 210061: contig of 36932 bp in length
 * 210062 210161: gap of 100 bp
 * 210162 227884: contig of 17723 bp in length.

FEATURES

source

1..227884
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-171P5"
 /clone.lib="RPI-23 Female Mouse BAC"

misc_feature

1..301
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left"

misc_feature

402..5831
 /note="assembly_fragment"

misc_feature

5932..173029
 /note="assembly_fragment"

misc_feature

173130..210061
 /note="assembly_fragment"

misc_feature

210162..227884
 /note="assembly_fragment"
 clone_end:17
 vector_side:right"

BASE COUNT 65784 a 48829 c 49902 g 62967 t 402 others
 ORIGIN

Query Match 69.6%; Score 445; DB 2; Length 227884;

Best Local Similarity 82.2%; Pred. No. 6.4e-52; Mismatches 100; Indels 15; Gaps 1;

Matches 530; Conservative 0; Mismatches 100; Indels 15; Gaps 1;

QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGCCACAGCTGAAGACACGGG 60
 |||||
 DB 87422 GAGGTGAACACTAAGCCAGCAGATCTCTAGTCTGAAGCCACAGCTGAAGATACCCGG 87481
 |||||
 QY 61 GCAAGCTGAGGGCTGAGGTGAGGACCCAGGACCTGGAGGGCCCTGCGCACCAG 120
 |||||
 DB 87482 GGCAGCTGATGAATGAGCTGAAGACACAGGACTTGGAGTGGCTCGGCCACCAA 87541
 |||||
 QY 121 GGCCTGAGCTGAGGTCTGTGAGATGATGCTGCAGCGCAAGAACGAGCGGAGCTG 180
 |||||
 DB 87542 GGGCTGAGCTGAGGTCTGTGAGAACCGAAGTCTGCAGCGCAAGAACGAGCAGACTG 87601
 |||||
 QY 181 CTGCGGAGAGAGTGAACCTGCTGGAGCAGGAGCTGAGGAGCTGGCGGCCAGCGCG -- 238
 |||||
 DB 87602 CTGAGAGAGAGTGNACCTCTAGACAGAGAGTGTGATGAGACTAGGGCCCGAGGCTGCT 87661
 |||||
 QY 239 -----CCTGCGCCCGCAGATGAGGCGGCCGCCACCTTCCCGAGGAGCTCCCT 285
 |||||
 DB 87662 CTGCATCCGGCACCCTCGGGGCTCCAGGGGTGGGCTCACTTTCTCAGAGGACATCCCT 87721
 |||||
 QY 286 GGCCTGCAGCGGAGCTGAGCGGCTCGGGCGGAGCTCGGGAGGAGCGGCAAGCCAT 345
 |||||
 DB 87722 GCTCTGACGGGAGCTGGATCGGCTCGGGCGGAGCTGAAGGAGGAGGCAAGGCCAT 87781
 |||||
 QY 346 GACCAGATCTCTCGGGCTTCAGCATGAGCGGCTGCTGTGAAGAGGAGAGAGGAGAG 405
 |||||
 DB 87782 GACCAGATCTCTCGGGATTCAGCATGAGCGGCTGCTGTGAAGAGGAGAGAGAGAA 87841
 |||||
 QY 406 GTGATTCAGTACAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCGCGGAACCCAG 465

Db 87842 GTGATTCAGTACCAAGGCAGCTGCAGCAGAGCTACTTTGGCCATGTACGCGCAACCAG 87901
 |||||
 QY 466 CGCTGGAGAGAGCCCTGCAGCAGCTGCAGCTGGGGACAGCGCGGGAGCCCTTTGGAG 525
 |||||
 DB 87902 GCTTAGAAGAGCCCTGCAGCAGCTGGCCGTGGGATGTTCACAGGAAACCCCTTGA 87961
 |||||
 QY 526 GTTGACCTGGAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGG 585
 |||||
 DB 87962 ATTGACCTGGAGGAGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGG 88021
 |||||
 QY 586 GCTGCTGGAGAGGGAGCTCTGGGACCTGGCACTGGGAGGGAGG 630
 |||||
 DB 88022 GCTGCTAGGGGTGGAGGGCTGGGACTGGACCCAGCAAGAGCAGG 88066
 |||||

RESULT 5

AC099416/c

LOCUS

AC099416 263546 bp DNA linear HTG 23-MAY-2002

DEFINITION

Mus musculus chromosome UNK clone RP23-122M11, WORKING DRAFT

SEQUENCE, 8 unordered pieces.

ACCESSION

AC099416

VERSION

AC099416.2 GI:21105058

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 263546)

McPherson,J.D. and Waterston,R.H.

Unpublished

JOURNAL

2 (bases 1 to 263546)

McPherson,J.D. and Waterston,R.H.

AUTHORS

Direct Submission

TITLE

Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE

3 (bases 1 to 263546)

McPherson,J.D. and Waterston,R.H.

AUTHORS

Direct Submission

TITLE

Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

JOURNAL

On May 23, 2002 this sequence version replaced gi:16924178.

COMMENT

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu

----- Project Information -----

Center project name: M.BA0122M11

----- Summary Statistics -----

Sequencing vector: M13; 32x

Sequencing vector: plasmid; 68x

Chemistry: Dye-primer ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.950319

Consensus quality: 278310 bases at least Q40

Consensus quality: 282938 bases at least Q30

Consensus quality: 286149 bases at least Q20

Insert size: 250000; agarose-fp

Insert size: 387570; sum-of-contigs

Quality coverage: 25.74 in Q20 bases; agarose-fp

Quality coverage: 16.86 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1204: contig of 1204 bp in length
 1205 1304: gap of unknown length
 1305 2524: contig of 1220 bp in length
 2525 2624: gap of unknown length
 2625 4136: contig of 1512 bp in length
 4137 4236: gap of unknown length
 4237 16434: contig of 12198 bp in length
 16435 16534: gap of unknown length
 32106 32107: contig of 15572 bp in length
 32107 32206: gap of unknown length
 32207 140224: contig of 108018 bp in length
 140225 140324: gap of unknown length
 140325 262933: contig of 122369 bp in length
 262934 262994: gap of unknown length
 262994 263546: contig of 753 bp in length.

FEATURES

source

1..263546
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosomes="UNK"
 /clone="RP23-122M11"
 1..1204
 /note="assembly_name:Contig127"
 1305..2524
 /note="assembly_name:Contig157"
 2625..4136
 /note="assembly_name:Contig360"
 4237..16434
 /note="assembly_name:Contig396"
 16535..32106
 /note="assembly_name:Contig397"
 32207..140224
 /note="assembly_name:Contig398"
 140325..262693
 /note="assembly_name:Contig399"
 262794..263546
 /note="assembly_name:Contig154"
 76628 a 56149 c 55623 g 74436 t 710 others

BASE COUNT

ORIGIN

Query Match 69.6%; Score 445; DB 2; Length 263546;
 Best Local Similarity 82.2%; Pred. No. 6.2e-52;
 Matches 530; Conservative 0; Mismatches 100; Indels 15; Gaps 1;

QY 1 GAGGTGAACCGAAGCTAGCAGATCTTGGGTCTCAAGGCACAGCTGAAGCAGCGGG 60
 DB 247464 GAGGTGAACCTAAGCCAGCAGATCTCTAGTCTGAAGGCACAGCTGAAGGATACCGGG 247405

QY 61 GCGAGCTGAGGGCTGAGCTGAGGCCAGGACCTGGAGGGCGCTGCGACCCAG 120
 DB 247404 GCGAGCTGAGTGAATGAGCTGAAGACAGGACTTGGAGGTGCTTGGAGGACCAAA 247345

QY 121 GCGCTGAGCTGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAAGCAGGCGGAGCTG 180
 DB 247344 GCGCTGAGCTGAGGTCTGTGAGACGAACTGCAGCGCAAGAAATGAGGACGACTG 247285

QY 181 CTGCGGGAAGTGAACTGCTGAGCAGGAGCTGCAGGAGCTGCGGGCCCGAGCGG-- 238
 DB 247284 CTGAGAGAGGTGAACCTTCTAGAACAGGAGTTGATGAGCTAGCGGCCCGAGGTGT 247225

QY 239 -----CCTGCGCCCGACATGGGGCGCCACCTCTCCCGAGGACGTCCCT 285
 DB 247224 CTGCATCCGGCACCCCTCGGGGCTCCAGGGGTTGGGCTCCTTCTCAGAGGACATCCCT 247165

QY 286 GCGCTGAGCGGAGCTGAGGCGGCTGCGGGCGGAGCTCGGGAGAGCGCAAGGCCAT 345
 DB 247164 GCTCTGACGGGAGCTGATCGGCTGCGGGCGGAGCTGAAGGAGGAGGCGCAAGGCCAT 247105

QY 346 GACCAGATGCTCGGGCTTCAGCATGAGCGGCTCGTGTGAAGGAGGAGGAGGAAG 405
 DB 247104 GACCAGATGCTCGGGATTCAGCATGAGCGGCTTGTGTGGAAGGAGGAGGAAG 247045

QY 406 GTGATTCACTACCAAGAACAGCTGCAGCAGAGCTAGTGGCCATGTACCAGCGGAACCAAG 465
 DB 247044 GTGATTCAGTACCAAGAGCAGCTGCAGCAGAGCTACTTGGCCATGTACCAGCGGAACCAAG 246985

QY 466 CGCCTGGAGAAAGGCCCTGCAGCAGCTGCAGCAGCTGGGGAAGCGCGGGGAGCCCTTGGAG 525
 DB 246984 CGTCTAGAGAAGGCCCTGCAGCAGCTGGCCGCTGGGATGTCTCCAGGGGAGCCCTTTGAA 246925

QY 526 GTTGACCTGGAAGGGGCTGACATCCCTACAGAGACATCATAGGCATCATGAGCTGAGCTGAGGG 585
 DB 246924 ATTGACCTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGGCATCATGAGCTGAGGG 246865

QY 586 GCTGCTGGGAAGCGAGCTCTGGGACCTGGCACTGGGAGGAGG 630
 DB 246864 GCTGCTAGGGGTGAGGGCTGGGACCTGGGACCTGGGAGGAGG 246820

RESULT 6

AC108987/c

LOCUS

DEFINITION

AC108987

AC108987.3

GI:21737647

HTG; HTGS PHASE1.

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 191210)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,

Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhaq,T., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Caron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlssohn,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Mosier,N., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okuwonu,G.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,

Scherer,S., Scott,G., Shen,H., Shoohrtari,N., Sisson,I.,

Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Umani,K., Vasquez,L., Veta,V., Villalob,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,

Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Neilson,D.,

Weinstock,G. and Gibbs.R.

Direct Submission

TITLE

JOURNAL

Unpublished

REFERENCE
AUTHORS

TITLE
JOURNAL

2 (bases 1 to 191210)
Worley,K.C.

Direct Submission
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS

TITLE
JOURNAL

3 (bases 1 to 191210)
Worley,K.C.

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 12, 2002 this sequence version replaced gi:18946600.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPMW
Center clone name: CH230-115K1
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115719 bases at least Q40
Consensus quality: 122885 bases at least Q30
Consensus quality: 129025 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 78 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1041: contig of 1041 bp in length
1042 1141: gap of unknown length
1142 2262: contig of 1121 bp in length
2263 2362: gap of unknown length
2363 3338: contig of 1176 bp in length
3339 3638: gap of unknown length
3639 4775: contig of 1137 bp in length
4776 4875: gap of unknown length
4876 6050: contig of 1175 bp in length
6051 6150: gap of unknown length
6151 7760: contig of 1510 bp in length
7761 9331: contig of 1571 bp in length
9332 9431: gap of unknown length
9432 11027: contig of 1596 bp in length
11028 11127: gap of unknown length
11128 12769: contig of 1642 bp in length
12770 12869: gap of unknown length
12870 14922: contig of 1223 bp in length
14923 14932: gap of unknown length
14933 15228: contig of 1036 bp in length
15229 15328: gap of unknown length
15329 16709: contig of 1381 bp in length
16710 16809: gap of unknown length
16810 18195: contig of 1386 bp in length
18196 18295: gap of unknown length
18296 19412: contig of 1117 bp in length
19413 19512: gap of unknown length
19513 20959: contig of 1447 bp in length
20960 21059: gap of unknown length
21060 22161: contig of 1102 bp in length
22162 22587: gap of unknown length
22588 23687: gap of unknown length

23688 25231: contig of 1544 bp in length
25232 25331: gap of unknown length
25332 27135: contig of 1804 bp in length
27136 27235: gap of unknown length
27236 29176: contig of 1940 bp in length
29177 30846: contig of 1571 bp in length
30847 30946: gap of unknown length
30947 32090: contig of 1144 bp in length
32091 32190: gap of unknown length
32191 34039: contig of 1849 bp in length
34040 34139: gap of unknown length
34140 35408: contig of 1269 bp in length
35409 35508: gap of unknown length
35509 37018: contig of 1510 bp in length
37019 37118: gap of unknown length
37120 38260: contig of 1142 bp in length
38261 38360: gap of unknown length
38362 40083: contig of 1723 bp in length
40084 40183: gap of unknown length
40184 41200: contig of 1017 bp in length
41201 41300: gap of unknown length
41302 43365: contig of 2065 bp in length
43366 43465: gap of unknown length
43467 45343: contig of 1878 bp in length
45344 45443: gap of unknown length
45444 46987: contig of 1544 bp in length
46988 47087: gap of unknown length
47088 49404: contig of 2317 bp in length
49405 49504: gap of unknown length
49506 50823: contig of 1319 bp in length
50824 50923: gap of unknown length
50924 52451: contig of 1528 bp in length
52452 52551: gap of unknown length
52552 54441: contig of 1890 bp in length
54442 54541: gap of unknown length
54543 56442: contig of 1901 bp in length
56443 56542: gap of unknown length
56543 57749: contig of 1207 bp in length
57750 57849: gap of unknown length
57850 60138: contig of 2289 bp in length
60139 60238: gap of unknown length
60239 62467: contig of 2229 bp in length
62468 62567: gap of unknown length
62568 64364: contig of 1797 bp in length
64365 64464: gap of unknown length
64465 66889: contig of 2425 bp in length
66890 66989: gap of unknown length
66990 68330: contig of 1841 bp in length
68331 68930: gap of unknown length
68931 71226: contig of 2296 bp in length
71227 71326: gap of unknown length
71327 73712: contig of 2386 bp in length
73713 73812: gap of unknown length
73813 76871: contig of 3059 bp in length
76872 76971: gap of unknown length
76972 79844: contig of 2873 bp in length
79845 79944: gap of unknown length
79945 81259: contig of 1315 bp in length
81260 81359: gap of unknown length
81360 83690: contig of 2331 bp in length
83691 83790: gap of unknown length
83791 86195: contig of 2405 bp in length
86196 86295: gap of unknown length
86296 88065: contig of 1770 bp in length
88066 88165: gap of unknown length
88166 90384: contig of 2219 bp in length
90385 90484: gap of unknown length
90485 91999: contig of 1515 bp in length
92000 92099: gap of unknown length
92100 94583: contig of 2484 bp in length
94584 94683: gap of unknown length
94684 98297: contig of 3614 bp in length

Query Match 68.6%; Score 438.6; DB 2; Length 191210;
Best Local Similarity 82.2%; Pred. No. 4.9e-51;
Matches 537; Conservative 0; Mismatches 94; Indels 22; Gaps 2;

QY 1 GAGGTGAACGCCAAGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 105989 GAGGTGAATCTTAGGCCAGTAGATCTTACCTTGAGGCACAGCTGAAGGACACCGG 105930

QY 61 GCCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCCCTTGCACCAAG 120
DB 105929 GCCAAGCTGGAGGAATGAGTTGAAGACACAGGACTTGGAGAGCGCTTGCACCAAG 105870

QY 121 GCGCTGGAGCTGGAGTCTGTGAGATGAGCTGCGAGCGCAAGAACGAGGCGAGCTG 180
DB 105869 GCGCTAGAGCTGGAGTCTGTGAGATGAAGCTGCGAGCGCAAGAAAGTGAAGGCGAGCTG 105810

QY 181 CTGCGGGGAGAGTGAACCTGCTGAGCAGGAGGCTGCGAGGCTGCGGGCCCGAGCGCC 240
DB 105809 CTGAGAGAGAGTGAACCTTTAGAACAGGAGCTGCTGGAGCTGCGGGCCCGAGCGCC 105750

QY 241 CTGCGGGCGGACATGGGGC-----GCCACCTTCCCGAGGAC 279
DB 105749 CTGCATCGAGATGCAGCACCCCTGGGCCCTCCAGGATTTGGGCTGACTTTCTCAGAGGAC 105690

QY 280 GTCCCTGCCCTCAGGGGAGCTGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCA 339
DB 105689 ATCCCTGTCTCAGCGGAGCTGAGTGGCTGCGGGCTGAGCTGAAGAGGAGAGGCA 105630

QY 340 GCCCATGACACAGATGCTCTCGGGCTTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAG 399
DB 105629 GCCCATGACACAGATGCTCTCGGGATTCAGCATGAGCGGTTGGTGTGAAGGAGGAG 105570

QY 400 GAGAGGTGAATCAGTACCAAGAACAGCTGCGAGGAGCTAGTGGCCACTGTACCAAGCGG 459
DB 105569 GAGAAAGTATCAGTACCAAGAACAGCTGCAACAGAGCTACTTGGCCATGTACCAAGCGC 105510

QY 460 AACGAGCGCTGAGAGGCGCTGAGCGGCTGCGAGCTGCGGAGGAGCGCGGGAGCGCC 519
DB 105509 AACGAGCGCTGAGAGGCGCTGAGCGGCTGCGGAGGAGCGCGGGAGGAGCGCC 105450

QY 520 TTGAGGTTGACTGCGAAGGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 579
DB 105449 TTTGAAATGACTGGAGGAGCTGACATTCCTATGAGGATATCATAGCCACTGAGATC 105390

QY 580 TGAGGGGTGCTT-GGGAAGGCGAGTCTGGGACCTTGGGACCTTGGAGGCGAGG 631
DB 105389 TGAGGGGTACCTCGGGGATGAGGCGCTGGGGACTAGACCCAGAAAGCAGG 105337

RESULT 7
AF123658
LOCUS Homo sapiens clone G3612 FE21 mRNA linear PRI 07-APR-1999
DEFINITION complete cds.
ACCESSION AF123658
VERSION AF123658.1 GI:4572473
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1722)
REFERENCE
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FE21 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.

TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES
source
1..1722
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="Bp22"
/clone="G3612"
1..1722
/gene="FE21"
1..1722
/gene="FE21"
/note="alternatively spliced"
/codon_start=1
/product="FE21"
/protein_id="A457247.1"
/db_xref="GI:4572474"
/translation="MGSVSLISGHSFKHCRASQYKLRKSSHLKLNRYSDGLLRF
GFSODSGHSGSSKMGSEDEFYIKVQKARGSHHPDYLTASSDGLGQAGVDFDPST
PKLMPFSNOLMGSEKGAVPFAPKPLPRSGALHSSPESASQHLHPAPDPKPEQ
ELKPLCSGALSDGRMSLPTHTSSSYQLDPLVTPVPTSRFGSSAHNITGCIIV
LQDNMMSLKALSFSDGSKLGHKNKADKGPSCVRSPTSTDECSTQLEOKLBRGA
LQKORSFEKELASSLAYERPRRDELEGPCKGNKQKQASQSQRAQQVHLHQ
VLQKQKQRQRQLESLMKEDLLETKLSYEREKTSFPALEETQMEVCQKGEIS
LLKQDKESQTEVNAKASEILGLKAQLKDTGRKLEGLRDTQDLEGLARTKGLLEVC
ENELQRKNAELLRKVNLEQLQELRAQAALARDMPPTTFEDVPALQRELERLV
WKEKEKVIQYQKQLQSYVAMYQRNQRLEKALQQLARGDSAGEPLEVLEGAIDPYE
DIATET"

BASE COUNT 381 a 541 c 563 g 237 t

ORIGIN
Query Match 67.9%; Score 434; DB 9; Length 1722;
Best Local Similarity 88.1%; Pred. No. 5.5e-50;
Matches 513; Conservative 0; Mismatches 0; Indels 69; Gaps 1;

QY 1 GAGGTGAACGCCAAGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 1210 GAGGTGAACGCCAAGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 1269

QY 61 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120
DB 1270 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 1329

QY 121 GGCCTGGAGCTGGAGGCTGTGAGAATGAGCTGAGCGCAAGAAAGAACGAGCGGAGCTG 180
DB 1330 GGCCTGGAGCTGGAGGCTGTGAGAATGAGCTGAGCGCAAGAAAGAACGAGCGGAGCTG 1389

QY 181 CTGCGGGGAGAGGTGAACCTCTGAGCAGGAGCTGAGGAGCTGCGGGCCCGAGCGCC 240
DB 1390 CTGCGGGGAGAGGTGAACCTCTGAGCAGGAGCTGAGGAGCTGCGGGCCCGAGCGCC 1449

QY 241 CTGCGCCCGACATGGGGCGCCCGCCACCTTCCCAGAGGAGCTCCCTGCGCTGAGCGGAG 300
DB 1450 CTGCGCCCGACATGGGGCGCCCGCCACCTTCCCAGAGGAGCTCCCTGCGCTGAGCGGAG 1509

QY 301 CTGAGAGCGGCTGCGGGCGCGAGCTGCGGGAGGAGCGGCAAGGCCATGACCATGTCTCTCG 360
DB 1510 CTG----- 1512

QY 361 GCGTTCACGATGAGCGGCTGTGTGAAGAGGAGGAGAGAGGAGGATTCAGTACCAAG 420
DB 1513 -----GAGCGGCTGTGTGAAGAGGAGGAGAGGAGGAGGATTCAGTACCAAG 1560

QY 421 AAACAGCTGACGAGAGCTAGTGGCCATGTACACGCGGAAACAGCGCTTGAGAGGCGC 480
DB 1561 AAACAGCTGACGAGAGCTAGTGGCCATGTACACGCGGAAACAGCGCTTGAGAGGCGC 1620

QY 481 CTGAGCAGCTGGCACTGTGGGAGCAGCGCCCGGGAGCGCTTGGAGGTTGACTGGAAGGG 540
DB 1621 CTGAGCAGCTGGCACTGTGGGAGCAGCGCCCGGGAGCGCTTGGAGGTTGACTGGAAGGG 1680

```
QY 541 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGA 582
|||||
Db 1681 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGA 1722
|||||

RESULT 8
AF123657
LOCUS Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
DEFINITION complete cds.
ACCESSION AF123657
VERSION AF123657.1 GI:4572471
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1692)
REFERENCE
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE
AUTHORS 2 (bases 1 to 1692)
Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
Direct Submission
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
JOURNAL Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES
source
1..1692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="G3611"
1..1692
/gene="FEZ1"
1..1692
/gene="FEZ1"
/translation="MGVSLSLISGHSFKHICRASQYKLRKSSHLKLNRYSDGLLRF
GFSQDSHGKSSSKSGEDFFYIKVSKARGSHHPDYTALSSDGLGGQAVDFDPST
PPKLPFPNQLMEGSEKGAIVRTAFKPYLPRSGAILHSSPESASHQLPAPDPKPEQ
ELKPLCSGALSDSGRNSMLPHTSTSSYQLOPLVTPVPTSRFGSGAHNIQTGIV
LQDSNMMSLKALSFSDGSKLGHSHKADKGPSCVRSPTSTDECSIQLEOKLREGA
LQKQRSFEELASSLAYERPRCRDELSGPEPKGNKLIKQASQKSQRAQVHLHQ
VLQKQKRQRLQBLSELMKQDLLETKLRSYERKTSFGPALETDQWECQKSGEIS
LQKQKESQTEEVNAKSEIILGLKAQLDKTGLGLELTDQLEGLARTKGLLEVC
ENELQKNEALREKYNLRLRAELRERQHQDQMSFGFQERLWVKEEKVQ
YQKQLQSYVAMYQNRQLEKALQQLARGDSAGEPLEVDLEGADIPYEDIITATSI"
BASE COUNT 380 a 522 c 554 g 236 t
ORIGIN

Query Match 58.5%; Score 374; DB 9; Length 1692;
Best Local Similarity 83.0%; Pred. No. 7.5e-42;
Matches 483; Conservative 0; Mismatches 0; Indels 99; Gaps 1;

QY 1 GAGGTGACGCCAGGCTAGCGAGTCTGGTCTCAAGGCACAGCTGAAGGACACGGG 60
|||||
Db 1210 GAGGTGACGCCAGGCTAGCGAGTCTGGTCTCAAGGCACAGCTGAAGGACACGGG 1269
|||||
QY 61 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120
|||||
Db 1270 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 1329
|||||
QY 121 GGCCTGGAGCTGGAGGCTGTGTGAGAAATGAGCTGCAGCGCGCAAGAACGAGGCGGAGCTG 180
|||||
```

Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 85530)
 Worley,K.C.
 Direct Submission
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 85530)
 Worley,K.C.
 Direct Submission
 Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GVR1
 Center clone name: CH230-72A23
 ----- Summary Statistics
 Sequencing vector: plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 3389 bases at least Q40
 Consensus quality: 35475 bases at least Q30
 Consensus quality: 36705 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 40 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1040: contig of 1040 bp in length
 * 1041: gap of unknown length
 * 1141: contig of 1322 bp in length
 * 2463: gap of unknown length
 * 2563: contig of 1013 bp in length
 * 3575: gap of unknown length
 * 3676: contig of 1014 bp in length
 * 4690: gap of unknown length
 * 4790: contig of 1434 bp in length
 * 6224: gap of unknown length
 * 6324: contig of 1149 bp in length
 * 7473: gap of unknown length
 * 7573: contig of 1533 bp in length
 * 9106: gap of unknown length
 * 9206: contig of 1673 bp in length
 * 10879: gap of unknown length
 * 10979: contig of 1449 bp in length
 * 12428: gap of unknown length
 * 12528: contig of 1460 bp in length
 * 13988: gap of unknown length
 * 14088: contig of 1077 bp in length
 * 15165: gap of unknown length
 * 15265: contig of 1017 bp in length
 * 16282: gap of unknown length
 * 16381: contig of 1679 bp in length
 * 18061: gap of unknown length
 * 18161: contig of 1979 bp in length
 * 20139: gap of unknown length
 * 20240: contig of 1085 bp in length
 * 21325: gap of unknown length
 * 21435: contig of 1809 bp in length
 * 23234: gap of unknown length
 * 23334: contig of 1718 bp in length
 * 25051: gap of unknown length
 * 25151: contig of 1523 bp in length
 * 26674: gap of unknown length
 * 26775: contig of 1560 bp in length
 * 28335: gap of unknown length
 * 28434: contig of 1411 bp in length
 * 29845: gap of unknown length
 * 29946: contig of 1215 bp in length
 * 31161: gap of unknown length
 * 31260: contig of 1054 bp in length
 * 32414: gap of unknown length
 * 32415: contig of 1737 bp in length
 * 34152: gap of unknown length
 * 34251: contig of 1480 bp in length
 * 34532: gap of unknown length
 * 35732: contig of 1195 bp in length
 * 35832: gap of unknown length
 * 37027: contig of 1764 bp in length
 * 37127: contig of 1855 bp in length
 * 38891: gap of unknown length
 * 38991: contig of 1575 bp in length
 * 40846: gap of unknown length
 * 40946: contig of 2032 bp in length
 * 42621: gap of unknown length
 * 44653: contig of 2327 bp in length
 * 47080: gap of unknown length
 * 47180: contig of 2345 bp in length
 * 49525: gap of unknown length
 * 49625: contig of 3145 bp in length
 * 52770: gap of unknown length
 * 52870: contig of 2552 bp in length
 * 55422: gap of unknown length
 * 55522: contig of 3245 bp in length
 * 58867: gap of unknown length
 * 58867: contig of 2523 bp in length
 * 61390: gap of unknown length
 * 61490: contig of 4036 bp in length
 * 65526: gap of unknown length
 * 65626: contig of 4439 bp in length
 * 70065: gap of unknown length
 * 70165: contig of 2926 bp in length
 * 73091: gap of unknown length
 * 73191: contig of 3872 bp in length
 * 77063: gap of unknown length
 * 77163: contig of 8368 bp in length.
 *
 * Location/Qualifiers
 * 1..85530
 * /organism="Rattus norvegicus"
 * /db_xref="taxon:10116"
 * /clone="CH230-72A23"
 *
 * BASE COUNT 21327 a 17928 c 18750 g 22171 t 5354 others
 * ORIGIN
 *
 * Query Match 36.4%; Score 232.4; DB 2; Length 85530;
 * Best Local Similarity 65.0%; Pred. No. 5.2e-23;
 * Matches 316; Conservative 0; Mismatches 166; Indels 4; Gaps 4;
 *
 * QY 148 GAGCTGCAGCGAAGAACGAGCGGAGGTGTGTCGGGAGAGGTGAACCTGCTGGAG 207

TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 1515)
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 2338 10th street, Philadelphia, PA 19107, USA
FEATURES
source
1. 1515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="D14"
1. 1515
/gene="FEZ1"
1. 1515
/gene="FEZ1"
/note="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="RAD23837.1"
/db_xref="GI:4572470"
/translation="MGVSLSLHSGHFKHCRASQYKLRKSHLKLNRVSDGLLRF
GFSODSGKSSKMGKEDFFYIKVQKAGSHHPDYALSSDGLGGQAGVDFDPT
PKLMPFNQLEMESEKAVPTAFKVLPRSGAILHSSPESASHQHPADPKPE
ELKPLCSGALSDSGRNSLSLPHSTSSVOLDPLVTPVGTFRFGSARNITQGV
LQDSNMWIKALSFGSKLHGNKADKSPVRSPISTDECSIQLEQLKLRREGA
LQKQRFEEELASLAYERPRRCDELGEPGPKGNKUKQASQKQACQVILHQ
VLQOQKQRLQLESLMKQDLLETKLRSYREKTSFGPALEETOWEVCQKGEIS
LQKQKQSETEVNAKASEILKQALDKLGRLELRTQDLGALRTGLGLEVC
ENELQSVVAMVQNRQLEKALQQLARGDSAGEPLEVDLEGADIPYEDIATEI"
BASE COUNT 336 a 483 c 482 g 214 t
ORIGIN
Query Match 27.4%; Score 175.4; DB 9; Length 1515;
Best Local Similarity 59.9%; Pred. No. 6.5e-15;
Matches 312; Conservative 0; Mismatches 206; Indels 3; Gaps 1;
QY 65 ACCTGAGGGCTGTGAGCTGAGGACCCAGGACCTGAGGGGCCCTGCGCACCAGGGCC 124
DB 995 ACCTGAGGTACTGACGTCTTACGAGGAGGAGCGGACGCTCGGAGGAGCTCGAGAGCC 1054
QY 125 TGGAGCTGGAGCTCTGTGAGATGAGCTGACGCGAAGAAACGAGCGGAGCTGCTGC 184
DB 1055 TCATGAGGAGCAGGACCTGCTGGAGACCAAGCTCAGTCTTACGAGGAGGAGACCA 1114
QY 185 GGGAGAGGTGAACCTGCTGGAGCAGGAGCTCAGGAGCTGCGGGCCCGCCCTGG 244
DB 1115 GCTTCGGCCCCGCTGGAGGAGCAGCCAGTGGAGGTGTGCGAAGTCCAGCGAGATCT 1174
QY 245 CCGCGACATGGGCGGCCCA---CCTTCCCGAGAGCTGCTCCTCCCTCGAGCGGGAGC 301
DB 1175 CCTCTCTGAAGCAGCAGCTGAAGGAGTCCAGACGAGGTGAACCGCAAGGCTAGCGAGA 1234
QY 302 TGGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCGG 361
DB 1235 TCTTGGTCTCAAGGCACAGCTGAGGACACGCGGGGCAAGCTGAGGGCCCTGGAGCTGA 1294
QY 362 GTTTCAGCATGAGCGGCTGTGTGGAAGGAGGAGAGAGTGATTCAGTACCAQA 421
DB 1295 GCACCCAGCACCTGGAGGCGGCCCTCGCACCAAGGGCCCTGAGCTGGAGGTCTGTGAGA 1354
QY 422 ACAGCTGACAGCAGCTACGTGGCCATCTACAGCGGAAACAGCGCCTGGAGAGGCC 481
DB 1355 ATGAGCTGACAGCAGCTACGTGGCCATCTACAGCGGAAACAGCGCCTGGAGAGGCC 1414
QY 482 TGCAGCAGCTGGCAGCTGGGAGCAGCGCGGGGAGCCCTTGGAGGTTGACCTGGAGGGG 541

Db 1415 TGCAGCAGCTGGCAGCTGGGGACAGCGCGGGGAGCCCTTGGAGTTGACCTGGAAGGGG 1474
QY 542 CTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 582
DB 1475 CTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1515
RESULT 12
BC005855
LOCUS Homo sapiens, clone IMAGE:2988487, mRNA, partial cds.
DEFINITION BC005855
ACCESSION BC005855.1 GI:13543390
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1612)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketteman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3 Row: n Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
FEATURES
source
1. 1612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clones="IMAGE:2988487"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC_15"
/lab_host="DH10B-R"
/note="vector: pOTB7"
<1. 934
/codon_start=2
/product="Unknown (protein for IMAGE:2988487)"
/protein_id="AAH05855.1"
/db_xref="GI:13543391"
/translation="ARGHWQREAREALREDCAAQRAQALQLQLVQLQKQKRSQ
LQDFALQREQLRRCATFLREELRWECVKQCKGEISLLKQKLESQ
AELVQKGSGLVALRVALRATLRVSEGRARGLQEARARELEACQSLQRHRE
AEQUREKAGOLDAAAGLRPPVPATADPFLAEDAKVORAAAGVGGSLRAQVER
LRLEQRRERREGEQDSFEGERLAWQAEKEQVIRYQKQLQHNLYIQMYRRNRQLESEL
QQLSLELEARELADLGLAEOAPCICLBETATEI"
BASE COUNT 348 a 458 c 539 g 267 t
ORIGIN
Query Match 23.2%; Score 148.2; DB 9; Length 1612;
Best Local Similarity 54.9%; Pred. No. 3.1e-11;
Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGTCTCAAGGCACACAGCTGAAGGACACGGCG 60
DB 311 GAGCTGGTGCAGAGGGCAGCGAGCTGGTGGTCTCTGCGGGTGGCGCTGCGGAGGCCCT 370

```
QY 61 GCACAGCTGGAGGCTTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 120
Db |||||
371 GCTACGCTGCGGTCAGTGAGGCGCTGCGGGGCTCTACAGGAGCGCGCCGAGCTCGG 430
QY 121 GGCCTGGAGCTGGAGGCTCTGAGATGAGCTGAGGCGCAAGAAAGAGGCGGAGCTG 180
Db |||||
431 GAGCTGGAGCTGGAGGCTGTTCCAGGAGCTGACGCGCACCCGCGCAGGAAGCTGAGCAG 490
QY 181 CTGCGGAGAGAGTGAACCTGCTGAGCAGGAGCTGACAGGAGCTGCGGCGC----- 230
Db |||||
491 CTGCGGAGAGAGCTGGGAGTGGATGCTGAGGCGCGGAGCTCGGGAGCGCCCTGTG 550
QY 231 -----CCAGGCGCGCTTGCGCGCGACATGGGCGCGCC 264
Db |||||
551 CCACCTGCCCGCTGACCACTTCTCTGGCAGAGAGTGATGAGGCCAAAGTGCGAGCGG 610
QY 265 ACCTTCCCGGAGAGCTGCTGCTGAGCGGAGCTGAGGCGCTGCGGGCGGAGCTG 324
Db |||||
611 GCAGCAGCGCGGTTGGGCGCAGCTTGCGGGCCAGGTGAGCGATTGCGGCTGAGCTG 670
QY 325 CGGAGGAGCGCAAGGCGCATGACAGATGCTCTCGGCTTCCAGCATGAGCGGCTGTG 384
Db |||||
671 CAGCGGAGCGCGCGGCTGAGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGCGC 730
QY 385 TGGAGGAGGAGAGGAGAGGTGATTGATTACCAAGAAACAGCTGCAGCAGCTACGCTG 444
Db |||||
731 TGGCAGGCGAGAGAGGAGCAGTGATCGCTACCAAGAGCAGCTGCAGCACAACCTACATC 790
QY 445 GCATGTACAGCGGAAACAGCGCTGAGAGGCGCTGAGAGGCGCTGCGAGCTGCGGAGAC 504
Db |||||
791 CAGATGTACCGCGCAACCGCAGCTAGACAGGAGCTGCAGCAGCTCAGCCTGAGCTG 850
QY 505 AGCGCGGCGGAGCTGAGGCTGACCTGG-----AAGGGCTGACATCCCTACGAG 558
Db |||||
851 GAGGCGCGGAGCTGCTGACCTGGGCTTGGCGGAGCAGCGCCCTGCTGATCTGCTGGAG 910
QY 559 GACATCATGACCTGAGATCTGAGGCGCTGCTGGAGGCGAGCTGCGGAGCTGCGCA 618
Db |||||
911 GAGATCACTGCTACTGAGATCTAGGCGCTCAGCAACAGCTCTCTAGGAGCTCTGCGCA 970
QY 619 CTGGG 623
Db |||
971 GAGGG 975
```

```
RESULT 13
HSM805394
LOCUS HSM805394 2275 bp mRNA linear PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp761K1711 (from clone DKFZp761K1711).
ACCESSION AL834338
VERSION AL834338.1 GI:21739969
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2275)
AUTHORS Anorge,W., Winkner,U., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp761K1711) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..2275
FEATURES
source
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="DKFZp761K1711"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2275
/gene="DKFZp761K1711"
<!-- 1597
/gene="DKFZp761K1711"
/notes="similarity to FEZ1 (Homo sapiens), different
splicing pattern"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD39005.1"
/db_xref="GI:21739970"
/translation="EKILIRPTAFKPVLPKPRGAPSLPSFMGRPATLGSQSGSLTQL
FGPASSSSSSSSSAADKPLAFSGWASGCPSTGLSDSRNLSLSLPTYSTGAEPTT
SSPGHLPSHSGRGALPGPARVPTGSHSDSGRSSKSTGLGRVAGLGSST
RASPDSSSCGERSPPPPPPDEALLHCVLEGLRDREAELQRLDLSDEATMCO
AYEPRQHWQREEREALEDCAQRAQRAQLLOLVFOLOQKROLODPFALLOE
REQURCATLREQRSLGPRLETKWEVCNGSEISLLKQQLKESQBELVQKSELV
ALRVAREARATLRVSEGRARGLQEAARELEACQELQRHREAEQJREKAGOL
DAEAAGREPPVPATADPFLAERDEAKVQAAAGVGSRLAVERLRVRELRERR
GEORDSFEGERLAWQAEQVIRYQKLOHNYIQWYRRNRQLBEQLQQLSLEARE
LADIGLAEQAPCILEEITATEI"
2236..2241
polyA_signal /gene="DKFZp761K1711"
2257
polyA_site /gene="DKFZp761K1711"
BASE COUNT 455 a 697 c 739 g 384 t
ORIGIN
Query Match 23.2%; Score 148.2; DB 9; Length 2275;
Best Local Similarity 54.9%; Pred. No. 2.9e-11;
Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGTCTCAAGGCACAGCTCAAGGACACGCGG 60
Db 974 GAGCTGTGTCAGAAAGGCGAGCTGTGTCTTGGGGTGGCGCTCGGAGGCGCCGT 1033
QY 61 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCAACCAAG 120
Db 1034 GCTACGCTGCGGGTCAGTGAGCGCGTCTACAGGAGCGCGCCGAGCTCGG 1093
QY 121 GGCCTGAGCTGGAGGCTGTGAGAAATGAGCTGAGCGCCAGAGAAAGAGCGGAGCTG 180
Db 1094 GAGCTGGAGCTGGAAGCCTGTTCCAGGAGCTGAGCGCACACCGCAGGAAGCTGAGCAG 1153
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGCAGCTGCGAGGAGCTGCGGCGC----- 230
Db 1154 CTGCGGAGAGAGGTGGGAGTGGAGTGTGAGCGCGCGGAGCTCGGAGCGCCCTGTG 1213
QY 231 -----CCAGGCGCGCCCTTGGCGCGGACATGGGCGCGCC 264
Db 1214 CCACCTGCCCGCTGAGCCATTCTCTGCGAGAGTGTGAGGCCAAAGTGCGAGCGG 1273
QY 265 ACCTTCCCGGAGGAGCTGCTGCTGCGCGGAGCTGGAGCGGCTCGGGCGCGAGCTG 324
Db 1274 GCAGCAGCGCGGTTGGGCGCAGCTTGGCGCCAGGTTGGAGCGATTTCGGGTGGAGCTG 1333
QY 325 CGGAGGAGCGGCAAGGCCATGACAGATGCTTGGGCTTCCAGCATGAGCGGCTGTG 384
Db 1334 CAGCGGAGCGCGCGGGGTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGCGGCTGGCC 1393
QY 385 TGGAGGAGGAGAGAGGAGAGGTGATTGATTACAGAAACAGCTGCAGCAGAGCTACGCTG 444
Db 1394 TGGCAGGAGAGAGGAGCAGGTGATCGCTTACAGAGAGAGCTGCAGCACAACCTACATC 1453
QY 445 GCCATGTACAGCGGAACCGCGCTGGAGAGGCGCTTGCAGCAGCTGCGGAGCTGGGAGAC 504
```


Db 1454 CAGATGTACCGCGCAACCGGAGCTAGACGAGGAGCTGCAGACCTCAGCTCGAGCTG 1513
 QY 505 AGCGCGGGGAGCCCTTGAGGTTGACCTGG-----AAGGGCTGACATCCCTACGAG 558
 Db 1514 GAGGCGCGGGAGCTCGCTGACCTGGGCTCGGCGGAGCGGCCCTGCATCTGCTGAG 1573
 QY 559 GACATCATGACCTGAGATCTGAGGGCTGCTCGGGAAGGCGAGCTTGGGACCTGCGCA 618
 Db 1574 GAGATCACTGCTACTGAGATCTAGGGCCCTAGCAACGAGCTCTGTAGGAGCTCTGCGCA 1633
 QY 619 CTGGG 623
 Db 1634 GAGGG 1638

RESULT 14
 BC006212 2766 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, clone MGC:2586 IMAGE:3161855, mRNA, complete cds.
 DEFINITION BC006212
 ACCESSION BC006212
 VERSION BC006212.1 GI:13623228
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2766)
 Direct Submission
 Submitted (09-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgs.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgs@nhgri.nih.gov

Shevchenko, Y., Weeteb, K.D., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,

Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,

Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,

Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: k Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction.

FEATURES
 source

1. .2766
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:2586 IMAGE:3161855"
 /tissue type="Brain, neuroblastoma"
 /clone lib="NIH MGC 19"
 /lab host="DH10B-R"
 /note="Vector: pOTB7"
 1138. .2088
 /codon_start=1
 /product="Unknown (protein for MGC:2586)"
 /protein_id="AAH06212.1"
 /db_xref="GI:13623229"
 /translation="MCQAEQRQHWREAREALREDCAAQRAQRAQLQLQVQFLQ

CDS

QOKROLODDPAOLLOBEROLERRCATLEREQRELGPRLBETKWEVCKGSEISLLKQ
 OLKESAEVYKGSSELVALREARPATRVSEGRARGLOEARAELEACSOEL
 QROQREARQERAGQDLDAEAGRPVPPATADPELLAESDEAKVQRAAGVGGSL
 RAQVRLURQLRRRREGRDSEFEGERLAWAEQEVIRYQKLOHNTIQMRRNR
 QLBQELQQLSLEARELADLGLAEAPACICLEITATEI"
 BASE COUNT 545 a 873 c 883 g 465 t

Query Match 23.2%; Score 148.2; DB 9; Length 2766;
 Best Local Similarity 54.9%; Pred. No. 2.8e-11;
 Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;

QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
 Db 1465 GAGCTGTGAGAGGCGGAGCTGGTGGCTCTGGGGTGGCGCTCGGGAGGCCCGT 1524
 QY 61 GGCAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGGCTCGGCACCAAG 120
 Db 1525 GCTACGCTGGGGTCACTGAGGGCGCTGCGGGGTCTACAGGAGGCGGCCGAGCTCGG 1584
 QY 121 GGCTTGGAGCTGGAGCTCTGTGAGATGAGCTGAGCGGCAAGAAAGACGAGCGGAGCTG 180
 Db 1585 GAGCTGGAGCTGGAAGCTGTTCACAGGAGCTGCAGCGACACCCGCCAGGAGCTGAGCAG 1644
 QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTGCGGCGC----- 230
 Db 1645 CTGCGGAGAGAGCTGGGAGTGGATGCTGAGCGGCGCGGACTCCGGGAGCCCTCTGTG 1704
 QY 231 -----CCAGCGCCCTTGGCGCCCGACATCGGGCGCGCC 264
 Db 1705 CCACCTGCCACCGCTGACCCCATTCCTCTGCGCAGAGTGTATGAGGCCAAAGTGCAGCGG 1764
 QY 265 ACCTTCCCGGAGGAGCTCCCTGCGCTGAGCGGAGCTGAGCGGCTGCGGGCGGAGCTG 324
 Db 1765 GCAGCAGCGGGGTGGGGGCGAGCTTGGGGCGCAGGTGGAGCGATTGCGGGTGGAGCTG 1824
 QY 325 CGGAGGAGGCGGCAAGCGCATGACAGATGCTCGGGCTTCACAGCATGAGCGGCTCGTG 384
 Db 1825 CAGCGGAGCGCGCGGGGTGAGGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGCGCC 1884
 QY 385 TGGAGGAGGAGAGAGGAGAGTGAATTGATTCAGTACAGAAACAGCTGCAGCAGAGCTACGTG 444
 Db 1885 TGGCAGGCGAGAGAGGAGCAGGTGATCCGTACCAAGACAGCTGCAGCACAACATACATC 1944
 QY 445 GCCATGTACACAGCGGACCGGCTGGAGAGGCGCTGAGAGGCGCTGCGAGCTGGCGGAC 504
 Db 1945 CAGATGTACCGCGCAACCGGCGAGCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTG 2004
 QY 505 AGCGCGGGGAGCGCTTGGAGGTTGACCTGG-----AAGGGGCTGACATCCCTTACGAG 558
 Db 2005 GAGGCCGGGAGCTGCTGACCTGGGCTGCGCAGCGAGCGGCCCTGCTGATCTGCTGAG 2064
 QY 559 GACATCATGACCTGAGATCTGAGGGGCTCGCTGGGAAGCGGAGTCTGGGGAGCTGCGCA 618
 Db 2065 GAGATCACTGCTACTGAGATCTAGGGCCCTCAGCAACCGAGCTCTGTAGGAGCTCTGCGCA 2124
 QY 619 CTGGG 623
 Db 2125 GAGGG 2129

RESULT 15
 AB058716 5733 bp mRNA linear PRI 05-JUN-2001
 LOCUS Homo sapiens mRNA for KIAA1813 protein, partial cds.
 DEFINITION AB058716
 ACCESSION AB058716
 VERSION AB058716.1 GI:14017842
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone:ph00819b.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)

AUTHORS

Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes. XX.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro
 DNA Res. 8 (2), 85-95 (2001)
 21245130
 2 (bases 1 to 5733)
 Direct Submission
 Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan [E-mail:cdmainfo@kazusa.or.jp,
 URL:htcjp://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
 Fax:81-438-52-3914]

FEATURES

Location/Qualifiers
 1..5733

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ph00819b"
 /tissue_type="brain"
 /note="vector:pBluescriptII SK plus"
 1..5733
 /gene="KIAA1813"
 /3052..5073
 /gene="KIAA1813"
 /note="Start codon is not identified."
 /codon_start=1
 /product="KIAA1813 protein"
 /protein_id="BAB47442.1"
 /db_xref="GI:14017843"

gene

CDS

/translation="ASATMAIVQTLVPLEPAEAATAPQPMGVSLSISGRPCPG
 GPAPRRHGPPTFRQDGLRGYEAQELCPAPVPRKAVPTFTYINEDFRTE
 SPSPSVDVDEQRANAHLPKPLIPVSGLEKNMEKILIRPTAFKPLPKPR
 GAPSPLPMGPATGLSGQSLTOLFGPSSSSSSSSSSAAADKPLAFSGWAGCPS
 GTLSDGNSLSLPTYTGTGAETTSPPGHLPSHSGRGALPGPARCVPTGPSHSD
 SGRSSSKTSGLGRVAGLNGSTRASPSSSCERSPPPPPPSPDEALLHCYLE
 GKLDRLEQLDLSNEATMCOAYEQRHWRERERAREALREDCAAQARAQAO
 LILOVFOQDLQDLFAQLQERLRCATLREQLRELPLEETKEVYCK
 SGEIILLQQLKSOAEIVKGSILVALRVALREARLTVSEGRGLQEAARREL
 ELEACQELORHQEAOLREKAGLDAAELREPPVPATDPLLAESDEAKVOR
 AAAGVGSIRLAQVELRVELORERREGRSEFGERLAWQAEKEQVIRYQKLOLHN
 YIQWYRRRLQELQQLSELEARELADLGLAEQAQPCILEEITATEI"

BASE COUNT

1048 a 1775 c 1815 g 1095 t

ORIGIN

Query Match 23.2%; Score 148.2; DB 9; Length 5733;
 Best Local Similarity 54.9%; Pred. No. 2.4e-11;
 Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;
 QY 1 GAGGTGAACGCCAGGCTAGCGAGATCTCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
 DB 4450 GAGCTGGTGCAGAGGCGAGCGAGCTGGTGGTCTCGCGGTGGCGCTGGCGGAGGCCGT 4509
 QY 61 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCGACCAAG 120
 DB 4510 GCTACGCTGCGGGTCAGTGAGGCGCGCTGGCGGGGTCTACAGAGGCGCGCGGAGCTCGG 4569
 QY 121 GGCCTGGAGCTGAGGCTCTGAGAAATGAGCTGACGCAAGCAAGAAAGAGCGGAGCTG 180
 DB 4570 GAGCTGGAGCTGGAAGCTTGTCCAGGAGCTGACAGGACACCGCCAGGAAGCTGAGCAG 4629
 QY 181 CTGCGGAGAGAGGTGAACCTGTGGAGGAGGAGCTGCAGGAGCTGCGGGC----- 230
 DB 4630 CTGCGGAGAGAGCTGGGCGAGTTGATGCTGAGGCGCGGAGCTCGGGAGCCCCCTGTG 4689
 QY 231 -----CMAGGCGCCCTGGCCCGCGGACATGGGGCGGCC 264
 DB 4690 CCACCTGCCACCGCTGACCCATTCTCTGGCGGAGAGTGATGAGGCCAAAGTGACGCGG 4749
 QY 265 ACCTTCCCGGAGGAGCTCCCTCCCTGCGAGCGGAGCTGGAGCGCTCGGGCCGAGCTG 324
 DB 4750 GCAGAGCCCGGGTGGGGGAGAGCTTGGCGGCGCCAGGTTGGAGCGATTGCGGGTGGAGCTG 4809

Search completed: June 15, 2003, 02:18:41
 Job time : 1667.94 secs

QY 325 CGGAGGAGCGGCAAGGCCATGACAGATGCTCTCGGGCTTCCAGCATGAGCGGCTCGTG 384
 DB 4810 CAGCGGAGCGCGCGCGGGTGAAGAGCAGCGGAGCAGCTTTTGAAGGGAGCGGCTGGCC 4869
 QY 385 TGGAGGAGGAGAGGAGAGAGGTGATTGAGTACCAAGAAACAGCTGCAGCAGAGCTACGTG 444
 DB 4870 TGGCAGGCGAGAGAGGAGAGGTGATCCGCTACCAAGAGCAGCTGCAGCAACTACATC 4929
 QY 445 GCATGTATACAGCGGAACCCAGCGCTGGAGAGGCCCTGCAGCAGCTGSCACGTGGGGAC 504
 DB 4930 CAGATGTATACCGCGCAACCGGAGCTAGAGAGGAGCTGCAGCAGCTAGCCTGGAGCTG 4989
 QY 505 AGCGCCCGGGAGCCCTTGGAGGTTGACCTGG-----AAGGGGCTGACATCCCTACGAG 558
 DB 4990 GAGGCGCGGAGCTCGTGACCTGGGCTGGCCGAGCAGGCCCCCTGCATCTGCTGGAG 5049
 QY 559 GACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAGGCGAGTCTGGGGACCTGGCA 618
 DB 5050 GAGATCACTGTACTGAGATCTAGGGCCCTCAGCAACCAAGCTCTGTAGGGAGCTCTGCGCA 5109
 QY 619 CTGGG 623
 DB 5110 GAGGG 5114

Result No.	Query	Score	Match	Length	DB	ID	Description
1	639	100.0	5492	21	AAA64508		cdna sequence of t
2	639	100.0	9048	21	AAA64507		Nucleotide sequenc
3	582	91.1	1791	21	AAA64509		cdna sequence enco
4	434	67.9	1722	21	AAA64615		Nucleotide sequenc
5	374	58.5	1692	21	AAA64614		Nucleotide sequenc
6	249	39.0	404	21	AAA64510		Nucleotide sequenc
7	224.4	35.1	1614	21	AAA64512		Nucleotide sequenc
8	172.4	27.0	1512	21	AAA64513		Nucleotide sequenc
9	149.2	23.3	1617	21	AAF15625		Human prostate can

PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
XX Example 2; Fig 5B; 255pp; English.
PS
PS The present sequence represents the cDNA sequence of the human FEZ1 gene.
XX FEZ1 is a tumour suppressor gene, located at chromosome location 8p22.
CC Decreased or no expression of FEZ1 is detected in a variety of cancer
CC cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
XX Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;
SQ
Query Match 100.0%; Score 639; DB 21; Length 5492;
Best Local Similarity 100.0%; Pred. No. 2.1e-99;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGTGAACGCCAAGGCTAGCAGATCTGGGTCTCAAGGCACAGCTGAAGACACGGG 60
DB 1321 GAGGTGAACGCCAAGGCTAGCAGATCTGGGTCTCAAGGCACAGCTGAAGACACGGG 1380
QY 61 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120
DB 1381 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 1440
QY 121 GCGCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAACGAGCGGAGCTG 180
DB 1441 GCGCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAACGAGCGGAGCTG 1500
QY 181 CTGCGGGAGAGGTGAACCTGTGGAGCAGGAGCTGCAGGAGCTCGGGCCCGCCAGCCGCG 240
DB 1501 CTGCGGGAGAGGTGAACCTGTGGAGCAGGAGCTGCAGGAGCTCGGGCCCGCCAGCCGCG 1560
QY 241 CTGCGCCCGCAGCATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGGAG 300
DB 1561 CTGCGCCCGCAGCATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGGAG 1620
QY 301 CTGGAGCGGCTCGGGCCGAGCTGCGGGAGGAGCGGCAAGCCATGACAGATGTCCTCG 360
DB 1621 CTGGAGCGGCTCGGGCCGAGCTGCGGGAGGAGCGGCAAGCCATGACAGATGTCCTCG 1680
QY 361 GCGTTCACGATGAGCGGCTCGTGGAGGAGGAGAGGAGAGAGGTGATTCAGTACCAG 420
DB 1681 GCGTTCACGATGAGCGGCTCGTGGAGGAGGAGAGGAGAGAGGTGATTCAGTACCAG 1740
QY 421 AAACAGCTGCAGCAGAGCTACGTGGCCATGTACACGCGGAACACGAGCGCTGGAGAGGCC 480
DB 1741 AAACAGCTGCAGCAGAGCTACGTGGCCATGTACACGCGGAACACGAGCGCTGGAGAGGCC 1800
QY 481 CTGACAGCTGCGACGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540
DB 1801 CTGACAGCTGCGACGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 1860
QY 541 GCTGACATCCCTAGCAGACATCATAGCCATGAGATCTGAGGGCTGCTGGGAAGGC 600
DB 1861 GCTGACATCCCTAGCAGACATCATAGCCATGAGATCTGAGGGCTGCTGGGAAGGC 1920
QY 601 GAGTCTGGGAGCTGGCACTGGGAGCGAGGGCTCTCCCG 639
DB

DB 1921 GAGTCTGGGAGCTGGCACTGGGAGCGAGGGCTCTCCCG 1959
RESULT 2
AAA64507
XX ID AAA64507 standard; DNA; 9048 BP.
XX AC AAA64507;
XX DT 02-JAN-2001 (first entry)
XX DE Nucleotide sequence comprising the human FEZ1 gene.
KW Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX Homo sapiens.
XX WO200050565-A2.
XX 31-AUG-2000.
XX 25-FEB-2000; 2000WO-US04950.
XX 25-FEB-1999; 99US-0121537.
XX (UVJE-) UNIV JEFFERSON THOMAS.
XX Croce CM, Ishii H;
XX WPI; 2000-558396/51.
XX New polynucleotide homologous with a portion of one strand of the human
FEZ1 gene, useful for alleviating abnormal cell proliferation such as
cancer -
XX Claim 2; Fig 5A; 255pp; English.
XX The present sequence comprises the human FEZ1 gene. FEZ1 is a tumour
suppressor gene, located at chromosome location 8p22. Decreased or no
expression of FEZ1 is detected in a variety of cancer cells. Expression
of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts
with tubulin, with microtubules, and with protein EPI-gamma.
XX Post-translational phosphorylation and dephosphorylation modulates the
effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are
useful for inducing cells to proliferate. Compounds which modulate FEZ1
association with tubulin are useful for alleviating tubulin hyper- or
hypo- polymerisation disorders, such as those associated with aberrant
initiation of mitosis, modulation of the initiation and rate of cell
proliferation and cell growth, modulation of cell shape, cell rigidity,
cell motility, rate and stage of cellular DNA replication, intracellular
distribution of organelles, metastatic potential of cell and cellular
transformation from a non-cancerous to cancerous phenotype. Compounds
which modulate FEZ1 binding and phosphorylation are also useful for
alleviating a disorder, such as tumorigenesis, tumour survival, growth
and metastasis.
XX Sequence 9048 BP; 2011 A; 2605 C; 2490 G; 1939 T; 3 other;
SQ
Query Match 100.0%; Score 639; DB 21; Length 9048;
Best Local Similarity 100.0%; Pred. No. 2.1e-99;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGTGAACGCCAAGGCTAGCAGATCTGGGTCTCAAGGCACAGCTGAAGACACGGG 60
DB 4912 GAGGTGAACGCCAAGGCTAGCAGATCTGGGTCTCAAGGCACAGCTGAAGACACGGG 4971
QY 61 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120
DB 4972 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 5031

QY 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAACGAGCGAGCTG 180
Db 5032 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAACGAGCGAGCTG 5091
QY 181 CTGCGGGAGAGGTGAACCTGCTGAGACAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 240
Db 5092 CTGCGGGAGAGGTGAACCTGCTGAGACAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 5151
QY 241 CTGGCCCGCAGCATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGAG 300
Db 5152 CTGGCCCGCAGCATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGAG 5211
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGCAAGGCCATGACACAGATGCTCTCG 360
Db 5212 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGCAAGGCCATGACACAGATGCTCTCG 5271
QY 361 GGCCTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGGAGAGAGCTGATTCAGTACCAG 420
Db 5272 GGCCTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGGAGAGAGCTGATTCAGTACCAG 5331
QY 421 AAACAGCTGCAGCAGCTGCTGCGCCATGTATACAGCGCAACCGCGCTGGAGAGGCC 480
Db 5332 AAACAGCTGCAGCAGCTGCTGCGCCATGTATACAGCGCAACCGCGCTGGAGAGGCC 5391
QY 481 CTGAGCAGCTGGCAGCTGGGAGACAGCGCGGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 540
Db 5392 CTGAGCAGCTGGCAGCTGGGAGACAGCGCGGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 5451
QY 541 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 600
Db 5452 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 5511
QY 601 GAGTCTGGGACCTGGCAGCTGGGAGCGAGGCGCTCTCCCG 639
Db 5512 GAGTCTGGGACCTGGCAGCTGGGAGCGAGGCGCTCTCCCG 5550

RESULT 3

AAA64509
ID AAA64509 standard; cDNA; 1791 BP.
XX
AC AAA64509;
XX
DT 02-JAN-2001 (first entry)
XX
DE cDNA sequence encoding a human FEZ1 polypeptide.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 1..1791
FT CDS /*tag= a
FT /product= "FEZ1"
XX
PN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UWJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Iehli H;
XX

DR WPI; 2000-558396/51.
DR P-PSDB; AAB08715.
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
XX cancer -
PS Claim 7; Fig 5I; 255pp; English.
XX
CC The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a
CC tumour suppressor gene, located at chromosome location 9p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;
Query Match 91.1%; Score 582; DB 21; Length 1791;
Best Local Similarity 100.0%; Pred. No. 8.6e-90;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTGGTCTCAAGGCACACAGCTGAAGCACGCGG 60
Db 1210 GAGGTGAACGCCAAGGCTAGCGAGATCTGGTCTCAAGGCACACAGCTGAAGCACGCGG 1269
QY 61 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120
Db 1270 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 1329
QY 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAACGAGCGAGCTG 180
Db 1330 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAACGAGCGAGCTG 1389
QY 181 CTGCGGAGAGAGTGAACCTGCTGAGCAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 240
Db 1390 CTGCGGAGAGAGTGAACCTGCTGAGCAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 1449
QY 241 CTGGCCCGCAGCATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCTGCAGCGGAG 300
Db 1450 CTGGCCCGCAGCATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCTGCAGCGGAG 1509
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 360
Db 1510 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 1569
QY 361 GGCCTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGAGAGAGCTGATTCAGTACCAG 420
Db 1570 GGCCTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGAGAGAGCTGATTCAGTACCAG 1629
QY 421 AAACAGCTGCAGCAGCTGCTGCGCCATGTATACAGCGGAGACCGCGCTGGAGAGGCC 480
Db 1630 AAACAGCTGCAGCAGCTGCTGCGCCATGTATACAGCGGAGACCGCGCTGGAGAGGCC 1689
QY 481 CTGAGCAGCTGGCAGCTGGGAGACAGCGCGGGAGAGCCCTTGGAGGTTGACCTGGAAGGG 540
Db 1690 CTGAGCAGCTGGCAGCTGGGAGACAGCGCGGGAGAGCCCTTGGAGGTTGACCTGGAAGGG 1749
QY 541 GCTGATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 582
Db 1750 GCTGATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791

```
RESULT 4
AAA64515
ID AAA64515 standard; cDNA; 1722 BP.
XX
AC AAA64515;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript G3612.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EF1-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..1722
FT /*tag= a
FT /product= "truncated FEZ1"
XX
PW WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Ishii H;
XX
DR WPI: 2000-558396/51.
DR P-PSDB; AAB08722.
XX
New polynucleotide homologous with a portion of one strand of the human
FEZ1 gene, useful for alleviating abnormal cell proliferation such as
cancer -
XX
PS Disclosure; Fig 5H; 255pp; English.
XX
The present sequence encodes a truncated human FEZ1 polypeptide. The
encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
suppressor gene, located at chromosome location 8p22. Decreased
or no expression of FEZ1 is detected in a variety of cancer cells.
Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
also interacts with tubulin, with microtubules, and with protein
EF1-gamma. Post-translational phosphorylation and dephosphorylation
modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
expression are useful for inducing cells to proliferate. Compounds
which modulate FEZ1 association with tubulin are useful for alleviating
tubulin hyper- or hypo- polymerisation disorders, such as those
associated with aberrant initiation of mitosis, modulation of the
initiation and rate of cell proliferation and cell growth, modulation of
cell shape, cell rigidity, cell motility, rate and stage of cellular
DNA replication, intracellular distribution of organelles, metastatic
potential of cell and cellular transformation from a non-cancerous to
cancerous phenotype. Compounds which modulate FEZ1 binding and
phosphorylation are also useful for alleviating a disorder, such as
tumorigenesis, tumour survival, growth and metastasis.
XX
Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;
SQ
Query Match 67.9%; Score 434; DB 21; Length 1722;
Best Local Similarity 88.1%; Pred. No. 8e-65;
Matches 513; Conservative 0; Mismatches 0; Indels 69; Gaps 1;
OY 1 GAGGTGAACGCCAAGCTAGCGAGATCTTGGGTCTCAAGGCACACAGCTGAAGGACACGCGG 60
```

PI Croce CM, Ishii H;
XX WPI; 2000-558396/51.
DR P-PSDB; AAB08721.
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
PS
PS Disclosure; Fig 5G; 255pp; English.
XX
XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EFl-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
XX Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;
SQ

Query Match 58.5%; Score 374; DB 21; Length 1692;
Best Local Similarity 83.0%; Pred No. 1.1e-54;
Matches 483; Conservative 0; Mismatches 0; Indels 99; Gaps 1;

QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACCTGAAGGACACGCGG 60
Db 1210 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACCTGAAGGACACGCGG 1269

QY 61 GCACAGCTGGAGGGCTGAGAGTGAAGACCTGAGGAGCGGCGCTTGGCGACCAAG 120
Db 1270 GCACAGCTGGAGGGCTGAGAGTGAAGACCTGAGGAGCGGCGCTTGGCGACCAAG 1329

QY 121 GSCCTGGAGCTGAGGCTCTGAGATGAGCTGACGCGCAAGAACGAGCGGAGCTG 180
Db 1330 GSCCTGGAGCTGAGGCTCTGAGATGAGCTGACGCGCAAGAACGAGCGGAGCTG 1389

QY 181 CTGCGGAGAGAGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGCGCCAGGCGCC 240
Db 1390 CTGCGGAGAGAGTGAACCT----- 1409

QY 241 CTGCGCGCGGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTCGCTGAGCGGGAG 300
Db 1410 -----G 1410

QY 301 CTGAGGCGGCTCGGGCGGAGCTCGGGAGGAGCGCAAGGCATGACAGATGTCCTCG 360
Db 1411 CTGAGGCGGCTCGGGCGGAGCTCGGGAGGAGCGCAAGGCATGACAGATGTCCTCG 1470

QY 361 GSCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGAGGTGATTCAGTACCAG 420
Db 1471 GSCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGAGGTGATTCAGTACCAG 1530

QY 421 AACAGCTGCACAGAGTACTGTGGCCATGTACACGCGGAACACGCGCTGAGAGGCG 480
Db 1531 AACAGCTGCACAGAGTACTGTGGCCATGTACACGCGGAACACGCGCTGAGAGGCG 1590

QY 481 CTGACAGAGCTGGACGCTGGGACACGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGG 540
Db 1591 CTGACAGAGCTGGACGCTGGGACACGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGG 1650

QY 541 GCTGACATCCCTACGAGGACATCATGACCCACTGAGATCTGA 582
Db 1651 GCTGACATCCCTACGAGGACATCATGACCCACTGAGATCTGA 1692

RESULT 6
AAA64510
ID AAA64510 standard; cDNA; 404 BP.
XX AC AAA64510;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript E16T6.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EFl-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..231
FT /tag= a
FT /product= "truncated FEZ1"
XX
PN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Ishii H;
XX
XX WPI; 2000-558396/51.
DR P-PSDB; AAB08717.
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
PS Disclosure; Fig 5C; 255pp; English.
XX
CC The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EFl-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;
Query Match 39.0%; Score 249; DB 21; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;

XX SQ Sequence 1617 BP; 355 A; 457 C; 532 G; 268 T; 5 other;
Query Match 23.3%; Score 149.2; DB 21; Length 1617;
Best Local Similarity 55.0%; Pred. No. 9.1e-17;
Matches 365; Conservative 0; Mismatches 258; Indels 41; Gaps 2;
QY 1 GAGGTGAACGCCAAGCTAGAGATCTGGGTCTCAAGGCACACGCTGAAGACACGCGG 60
DB 279 GAGCTGGTGCAGAGGCGAGCTGGTGGCTCTGCGGGTGGCTCGGAGGCCCCGT 338
QY 61 GGCAAGCTGGAGGCGCTGAGCTGAGGACCTGAGGCGCCCTCGGCACCAAG 120
DB 339 GCTACGCTCGGGTCACTGAGGCGCTGCGGGGTCTACAGAGGCGCCCGAGCTCG 398
QY 121 GGCCTGGAGCTGAGGCTGTGAGATGAGCTGAGCGCAAGAAAGAGCGGAGCTG 180
DB 399 GAGCTGGAGCTGGAAGCCCTGTTCCAGGAGCTGACAGCGACACCGCCAGGAAGCTGAGCAG 458
QY 181 CTGCGGAGAGGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTGCGGGC----- 230
DB 459 CTCGCGGAGAAAGCTGGGCGAGTTGATGCTGAGCGCGCGGACTCCGGAGCCCCCTGTG 518
QY 231 -----CCAGGCGCGCCCTGGCCCGGACATGGGGCGGCC 264
DB 519 CCACCTGCCACCGCTGACCCATTCTCTGGCAGAGAGTGATGAGGCCAAGTGACGCGG 578
QY 265 ACCTTCCCGAGGAGCTCCCTGCTGCGAGCGGGAGCTGGAGCGCTCGGGCCGAGCTG 324
DB 579 GCAGCAGCGGGGTTGGGGGCGAGCTTGGGGCCGAGGTGGAGCGATTGCGGGTGGAGCTG 638
QY 325 CGGAGGAGCGGAGGCGCATCAGAGTCTCTCGGGCTCCAGCATGAGCGGCTCGTG 384
DB 639 CAGCGGAGCGGCGGGGTGAGGAGCAGCGGACAGCTTGGAGGGGAGCGGCTGGCC 698
QY 385 TGAAGAGGAGGAGAGAGAGTGTATTGATTACAGAAACAGCTGCAGCAGAGCTTACCTG 444
DB 699 TGGCAGGCGAGAGAGAGCAGGTGATCCGCTACAGAGAGCAGCTGCAGCAGACTACATC 758
QY 445 GCATGTATCAGCGGAACAGCGCTCGAGAGGCGCTGAGAGCGCTGCAGCTGCGAGTGGGAC 504
DB 759 CAGATGTATCCGCGCAACCGGAGCTAGAGCAGGAGCTGCAGCAGCTCAGCTGAGCTG 818
QY 505 ACGCGCGGGAGCCCTTGAGGTTGACCTG-----GAAGGGCTGCACATCCCTACGAGG 559
DB 819 GAGGCGCGGAGCTGCTGACCTGGGCTTGGCGGAGAGCCCTGCTATCTCCCTGGAGG 878
QY 560 ACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGCGAGTCTGGGACCTGGCAC 619
DB 879 AGATCACTGCTACTGAGATCTAGGGCCCTCAGCAACCCAGCTCTGTAGGAGCTCTGCCAG 938
QY 620 TGGG 623
DB 939 AGGG 942

RESULT 10
ABQ54361
ID ABQ54361 standard; cDNA; 1623 BP.
XX AC ABQ54361;
XX AC
DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HELGN26 cDNA, SEQ ID NO:241.
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.
OS Homo sapiens.
PN WO200200677-A1.
PD 03-JAN-2002.
XX 07-JUN-2001; 2001WO-US18569.
PF 07-JUN-2000; 2000US-209467P.
PR (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
DR P-PSDB; ABP41284.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX Claim 1; SEQ ID No 241; 2922pp; English.
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54361-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovarian and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1623 BP; 358 A; 457 C; 534 G; 268 T; 6 other;
Query Match 23.2%; Score 148.2; DB 24; Length 1623;
Best Local Similarity 54.9%; Pred. No. 1.3e-16;
Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACACGCTGAAGACACGCGG 60
DB 282 GAGCTGGTGCAGAGGCGGCGAGCTGGTGGCTCTGCGGGTGGCTCGGAGGCCCCGT 341
QY 61 GGCAGCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGGAGGGGCGCCCTCGGCACCAAG 120
DB 342 GCTACGCTGCGGGTCACTGAGGCGCGCTGCGCGGGTCTACAGGAGGCGCGCGGAGCTCGG 401

PT Composition for detecting protein-protein interactions in a mammalian
PT two hybrid system comprises bait and test vector which consist of
PT selection gene, vector viral origin of replication and fusion gene -
XX
XX Disclosure; Column 17-20; 18pp; English.
XX
CC The present sequence represents the Epstein Barr nuclear antigen.
CC It is used to produce bait vectors of the invention. The specification
CC describes a compositions and methods for a genetic system of detecting
CC protein-protein interactions in a mammalian host cell. The system
CC comprises bait and test, both containing selection genes, and viral
CC origin of replications which require bound viral replication proteins
CC to effect replication. The compositions is useful for detecting an
CC interaction between a bait protein and a test protein. It is useful in
CC a mammalian two-hybrid system for detecting protein-protein interactions
CC in a mammalian host cell.
XX
SQ Sequence 2580 BP; 632 A; 512 C; 1054 G; 382 T; 0 other;

Query Match	17.9%	Score 114.4;	DB 21;	Length 2580;
Best Local Similarity	49.8%	Pred. NO. 6.8e-11;		
Matches 317; Conservative	0;	Mismatches 316;	Indels. 3;	Gaps 1;

QY	1	GAGGTGAACCCCAAGGCTACGGAGTCTCTGGGTCTCAAGGCA	CAGCTGAAGGACACGCCG	60
Db	847	GAGGGGCAGCAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGG		906
QY	61	GGCAAGCTGAGGGCCCTGGAGCTGAGACCCAGGACCTGAGGGCGCCCTGCCACCCAG		120
Db	907	CAGGAGCAGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAG		966
QY	121	GGCCTGAGCTGGAGGCTCTGTGAAATGAGCTGCAGCGCAAGAAAGAACGAGGGCGGAGCTG		180
Db	967	GGCAGGAGCAGGAGGAGGGGCAGGAGCCAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAG		1026
QY	181	CTCGGGGAGAGGTGAACCTGCTGAGCAGCAGGAGCTGCAGAGAGTGCGGGGCCGAGGCCGCC		240
Db	1027	CAGGAGGGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG		1086
QY	241	CTGGCCCGCACATGGGGCCGCCACCTTCCCGAGGACCTCTGCTGCCCTGCAGCGGGAG		300
Db	1087	CAGGAGCAGGAGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAG		1146
QY	301	CTGGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGGGCAAGGCCATGACAGATGTCTCTCG		360
Db	1147	CAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG		1206
QY	361	GGCTTCAGCATCAGCGGCTCGTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG		420
Db	1207	CAGGAGGGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG		1266
QY	421	AAACAGCTGCAGCAGAGCTACGTGGCCATGTACCGCGAACCCAGCGCCCTGGAGAGGCC		480
Db	1267	CAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGGAGGAGG		1325
QY	481	CTGCAGCAGCTGGCAGCTGGGGACAGCCCGGGGAGCCCTTGAGAGTTGACTTGGAGGGG		540
Db	1326	--GGAGGGGCAGAGCAGGAGGGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG		1383
QY	541	GCTGACATCCCTACAGGAGCATCATAGCCACCTGAGATCTGAGGGGCTGCTGGGAGGC		600
Db	1384	GTGAGGCCGGGGTCGAGGAGGTAGTGGAGGCCGGGGTCGAGGAGGTAGTGGAGGCCGCC		1443
QY	601	GAGTCTGGGACCTGGCACTGGGAGGAGGGGCTCTC		636
Db	1444	GGGGTGAAGACCTGAAAGACCAAGGGGGGGAAGTC		1479

RESULT 14
AAI64275
ID AAI64275 standard; DNA; 2580 BP.
XX
AC AAI64275;

XX	14-MAR-2002	(first entry)
DT		
XX	Epstein-Barr virus nuclear antigen coding sequence.	
DE		
XX		
KW	Epstein barr virus; protein-protein interaction; fusion protein;	
KW	yeast two-hybrid; bioactive agent; therapeutic drug; ds.	
XX		
OS	Human herpesvirus type 4.	
XX		
PN	US6316223-B1.	
XX		
PD	13-NOV-2001.	
XX		
PF	22-JUL-1999; 99US-0359081.	
XX		
PR	30-MAR-1998; 98US-0050863.	
XX		
PA	(RIGE-) RIGEL PHARM INC.	
XX		
PI	Payan D, Luo Y, Huang B;	
XX		
DR	WPI; 2002-088877/12.	
XX		
PT	Detecting protein-protein interactions, useful screening candidate	
PT	bioactive agents or therapeutic drugs, comprises employing mammalian	
PT	cloning system cell comprising a bait vector and a test vector -	
XX		
PS	Disclosure; Column 27-30; 29pp; English.	
XX		
CC	The invention relates to the detection of an interaction between a bait	
CC	protein and a test protein, comprising employing a mammalian host cell	
CC	comprising a bait vector and a test vector, where the vectors contain	
CC	genes that encode proteins, and the method uses fusion proteins whose	
CC	expression and interaction with each other results in transcriptional	
CC	activation. The method and vectors are useful for analysing	
CC	protein-protein interactions between known proteins and for isolating,	
CC	cloning or characterising unknown proteins. The method is also useful	
CC	for screening candidate bioactive agents that modulate the	
CC	protein-protein interaction between a bait protein and a test protein	
CC	and thus are useful for the identification of novel therapeutic drugs.	
CC	The present method employs mammalian cloning systems that are highly	
CC	stable, and designed to reduce the background signals frequently found	
CC	in other systems such as the yeast two-hybrid system. The current	
CC	sequence represents the coding sequence for the Epstein-Barr virus	
CC	nuclear antigen, the viral replication protein, which, on binding to the	
CC	Epstein-Barr origin of replication (see AAI64274) on the bait vector,	
CC	allows vector replication to occur	

Sequence 2580 BP; 632 A; 512 C; 1055 G; 381 T; 0 other;

Query Match 17.9%; Score 114.4; DB 24; Length 2580;
Best Local Similarity 49.8%; Pred. No. 6.8e-11;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

Qy	1	GAGGTGAACCCCAAGGCTAGCGATGCTTGGGTCTCAAGGCACAGCTGTAAGGACACGCGG	60
Db	847	GAGGGCAGAGCAGGAGGAGGGGCGCAGGAGGGGCAAGAGGGGCAGGAGCAGGAGGAGGGG	906
Qy	61	GGCAAGCTGGAGGGCCTGGAGCTTGAGACCCAGGACCTTGAGAGGCGCCCTGCGCACCAAG	120
Db	907	CAGGAGCAGGAGGAGGGCAGGAGGGGCGCAGGAGCAGGAGGAGGGCAGGAGGGGCGAGG	966
Qy	121	GGCTCGAGCTGAGGTCTGTGAGATGAGTCCAGCGCAAGAGAACGACGAGGCGGAGCTG	180
Db	967	GGCGAGGACAGGAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCGAGGAG	1026
Qy	181	CTCCGGGAGAGGTGAACCTGCTGGAGCAGGAGGCTGCAGGAGCTGCGGGGCCAGGCGGCC	240
Db	1027	CAGGAGGGGCAGGAGCAGGAGGAGGGGCGCAGGAGGGGCAGGAGGGGCAGGAGGAGGGG	1086
Qy	241	CTGGCCCGCGACATGGGGCGCGCCACCTTCCCGAGGACGCTCCCTGCTCCCTCAGCGGGAG	300

Db 1087 CAGGAGCAGGAGGCGCAGGAGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAG 1146
QY 301 CTGGAGCGGCTCGGCGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACAGATGTCCTCG 360
Db 1147 CAGGAGGCGCAGGAGGCGCAGGAGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAG 1206
QY 361 GCCTCCAGCATGAGGCGCTGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 1207 CAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAG 1266
QY 421 AAACAGCTGCACAGCATGCTACCTGCGCCATGTACACGCGGAACAGCGCTGGAGAAGGCC 480
Db 1267 CAGGAGCAGGAGGCGCAGGAGGCGCAGGAGCAGGAGGCGCAGGAGGCGCAGGAGCAGGA- 1325
QY 481 CTGACAGCTGCGCAGCTGGGACAGCGCCGGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 540
Db 1326 --GGAGGCGCAGGAGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAGGCGCAGGAGGCA 1383
QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 600
Db 1384 GTGAGGCGCGGGTCTGAGGAGGTAGTGGAGGCGCGGGTCTGAGGAGGTAGTGGAGGCGCGCC 1443
QY 601 GAGTCTGGGACCTGGCACTGGGAGGCGAGGCGCTCTC 636
Db 1444 GGGGTAGAGGACGTGAAGAGCCAGGCGGGGGAAGTC 1479

RESULT 15

AAx90923/c

ID AAX90923 standard; DNA; 5452 BP.

XX

AC AAX90923;

XX

DT 17-JAN-2000 (first entry)

XX

DE Anti-sense strand of pCMVBENA plasmid.

XX

KW Anti-sense strand; plasmid pCMVBENA; EBNA 1; episome; transfection;

KW Epstein Barr Virus Nuclear Antigen 1; origin of replication;

KW EBV oriP; eukaryotic host cell; recombinant cell line; ion channel;

KW multiple gene expression; receptor; transporter protein; gene therapy;

KW transcription factor; adhesion molecule; antisense therapy;

KW gene amplification; cell immortalisation; ss.

XX

OS Epstein-barr virus.

OS Cytomegalovirus.

OS Synthetic.

XX

FH Location/Qualifiers

FT CDS

FT complement (3032..4957)

FT /*tag= a

FT /product= "EBNA 1 protein"

FT /note= "Epstein Barr Virus Nuclear Antigen 1"

XX

PN W09947647-Al.

XX

PD 23-SEP-1999.

XX

PF 12-FEB-1999; 99WO-US03307.

XX

PR 18-MAR-1998; 98US-0040961.

XX

PR 06-AUG-1998; 98US-0130114.

XX

XX (PHAR-) PHARMACOEPIA INC.

XX

PI DamaJ BB, Horlick RA, Robbins AK;

XX

XX WPI; 1999-610610/52.

XX

XX New method for expressing genes from recombinant eukaryotic cells,

XX useful for gene therapy

XX

XX Example 1; Fig 1; 86pp; English.

XX

CC The present sequence is an anti-sense strand of commercially available
CC plasmid pCMVBENA from which Epstein Barr Virus Nuclear Antigen 1 (EBNA
CC 1) encoding DNA is obtained. EBNA 1 protein is used to stably maintain
CC episomes containing EBV origin of replication (oriP) and a gene encoding
CC protein or RNA of interest. Eucaryotic host cells expressing EBNA 1
CC protein are transfected with these episomes to produce recombinant
CC cell lines expressing multiple genes of interest. This provides a
CC rapid and reliable method of stably expressing multiple genes in
CC transfected cells. The episomes are useful in the transfection of genes
CC encoding receptors, transporter proteins, ion channels, adhesion
CC molecules and transcription factors. The episomes carrying desired genes
CC can also be used to transfect cells in gene therapy, antisense therapy,
CC for gene amplification, cell immortalisation, etc.

SQ Sequence 5452 BP; 1108 A; 1736 C; 1246 G; 1362 T; 0 other;

Query Match

Best Local Similarity 17.9%; Score 114.4; DB 20; Length 5452;

Matches 317; Conservative 49.8%; Pred. No. 6.8e-11;

Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACGCAAGGCTAGCGAGATCTGGTCTCAAGGCACACCTGAAGGACACGCGG 60

Db 1958 GAGGGCAGGAGCAGGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGAGGCGG 1899

QY 61 GGCAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCGCTCGCCACCAAG 120

Db 1898 CAGGAGCAGGAGGCGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1839

QY 121 GGCCTGGAGCTGGAGGCTCTGTGAGAACTGAGCTGAGCGCAAGAAAGCGCGGAGCTG 180

Db 1838 GGGCAGGAGCAGGAGGCGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1779

QY 181 CTGCGGGAAGAGGTGAACCTCTCTGAGCAGGAGCTGCGAGGAGCTGCGGGCCAGCGCCG 240

Db 1778 CAGGAGGCGGCGAGGAGCAGGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1719

QY 241 CTGCGCCCGACATGGGCGCCCGCCACCTTCCCGAGGAGCTCCCTGCGCTGCGAGCGGAG 300

Db 1718 CAGGAGCAGGAGGCGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1659

QY 301 CTGAGCGCGCTGCGGGCGCGAGCTGCGGAGGAGCGGCAAGGCCATGACAGATGCTCTCG 360

Db 1658 CAGGAGGCGCAGGAGGCGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGAGG 1599

QY 361 GGTTCAGCATGAGCGGCTCTGTGGAAGGAGGAGAGAGAGAGTTCAGTACCAAG 420

Db 1598 CAGGAGGCGGCGAGGAGCAGGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1539

QY 421 AAACAGCTGCACAGCATGCTAGTGGCCATGTACCAAGCGGCAAGCCAGCGCTGGAGAAGCC 480

Db 1538 CAGGAGCAGGAGGCGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGGCGGCGAGG 1480

QY 481 CTGACAGCTGGCACAGTGGGCGACAGCGCCCGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 540

Db 1479 --GGAGGCGCAGGAGCAGGAGGCGGCGGCGAGGAGGCTGGAGGCGCGGGTCTGAGGAGCA 1422

QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGCTCGGAAGGC 600

Db 1421 GTGAGGCGCGGGTCTGAGGAGGTAGTGGAGGCGCGGGTCTGAGGAGGTAGTGGAGGCGCGCC 1362

QY 601 GAGTCTGGGACCTGGCACTGGGAGGCGAGGCTCTC 636

Db 1361 GGGGTAGAGGACGTGAAGAGCCAGGCGGGGGAAGTC 1326

Search completed: June 14, 2003, 23:43:06

Job time : 149.118 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	114.4	17.9	1926	4	US-09-249-585A-2	Sequence 2, Appli	
2	114.4	17.9	2580	3	US-09-050-863-2	Sequence 2, Appli	
3	114.4	17.9	2580	4	US-09-359-081-2	Sequence 2, Appli	
C	4	114.4	17.9	5452	2	US-09-130-114-1	Sequence 1, Appli
	5	114.4	17.9	9600	4	US-08-910-647-1	Sequence 1, Appli
6	114.4	17.9	9600	4	US-09-620-925-1	Sequence 1, Appli	
7	114.4	17.9	10596	1	US-07-884-811-15	Sequence 15, Appl	
8	114.4	17.9	10596	1	US-07-885-971-15	Sequence 15, Appl	
9	114.4	17.9	10596	1	US-08-087-783A-15	Sequence 15, Appl	
10	114.4	17.9	10596	1	US-08-194-088B-15	Sequence 15, Appl	
11	114.4	17.9	10596	2	US-08-194-087-15	Sequence 15, Appl	
12	114.4	17.9	10596	5	PCR-US93-04648-15	Sequence 15, Appl	
13	88.8	13.9	3489	2	US-08-728-323A-1	Sequence 1, Appli	
14	88.8	13.9	3489	4	US-09-298-568-1	Sequence 1, Appli	
C	15	88.8	13.9	32207	2	US-08-770-379-20	Sequence 20, Appl
	16	88.8	13.9	32207	4	US-08-757-669A-20	Sequence 20, Appl
C	17	88.8	13.9	32207	4	US-08-730-371A-20	Sequence 20, Appl
	18	88.6	13.9	2793	1	US-08-209-747-1	Sequence 1, Appli
19	88.6	13.9	2793	1	US-08-458-298-1	Sequence 1, Appli	
C	20	81.2	12.7	1926	4	US-09-249-585A-4	Sequence 4, Appli
	21	81.2	12.7	1931	2	US-09-130-114-2	Sequence 2, Appli
C	22	81.2	12.7	16442	3	US-08-781-891-208	Sequence 208, App
	23	78.4	12.3	9551	1	US-08-056-200-93	Sequence 93, Appl
24	78.4	12.3	9551	2	US-08-800-644-93	Sequence 93, Appl	
25	76	11.9	2338	1	US-08-425-069-1	Sequence 1, Appli	
26	76	11.9	2338	2	US-08-317-844B-1	Sequence 1, Appli	
27	75	11.7	1995	1	US-08-425-069-3	Sequence 3, Appli	

REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match 17.9%; Score 114.4; DB 4; Length 2580;
Best Local Similarity 49.8%; Pred. No. 4.7e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACGCCAAGCTAGCAGATCTCTGGGTCTCAAGGCACACCTGAAGACACGCGG 60
DB 847 GAGGGGACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
QY 61 GSCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAG 120
DB 907 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
QY 121 GSCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGCGACGCAAGAAAGAGGCGGAGCTG 180
DB 967 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGCGAGGAGCTGCGGGCCAGGCGGCC 240
DB 1027 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
QY 241 CTGGCCCGACATGGGGCCGCCACCTTCCCGAGAGCTCCCTGCGCCCTGAGCGGGAG 300
DB 1087 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 1147 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
QY 361 GCCTTCAGCATGACGGCTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 1207 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
QY 421 AAACAGCTGCACAGCTACTGCGCCATGTACACGCGAACCCAGGCGCTTGAGAGAGGCC 480
DB 1267 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1325
QY 481 CTGACAGCTGGCCACGCTGGGACACAGCGCCGGGAGCCCTTGAGGCTTGACCTGGAAGG 540
DB 1326 --GGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1383
QY 541 GCTGACATCCCTACAGAGACATCATAGCCATGAGATCTGAGGGGCTGCTTGGGAAGGC 600
DB 1384 GTGAGGCGGGGCTCGAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1443
QY 601 GAGTCTGGGACCTGCGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
DB 1444 GGGGTAGAGGAGCTGAAGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1479

RESULT 4
US-09-130-114-1/c
Sequence 1, Application US/09130114
Patent No. 597807
GENERAL INFORMATION:
APPLICANT: Horlick, Robert A.
APPLICANT: Damaj, Basam B.
APPLICANT: Robbins, Alan K.
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
FROM MULTIPLE TRANSFECTED EPISOMES

FILE REFERENCE: 0867/1D903US1
CURRENT APPLICATION NUMBER: US/09/130.114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PastSeq for Windows Version 3.0
SEQ ID NO: 1
LENGTH: 5452
TYPE: DNA
ORGANISM: VEBNA
US-09-130-114-1

Query Match 17.9%; Score 114.4; DB 2; Length 5452;
Best Local Similarity 49.8%; Pred. No. 5.2e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACGCCAAGCTAGCAGATCTCTGGGTCTCAAGGCACACCTGAAGACACGCGG 60
DB 1958 GAGGGGACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1899
QY 61 GSCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAG 120
DB 1898 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1839
QY 121 GSCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGCGACGCAAGAAAGAGGCGGAGCTG 180
DB 1838 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1779
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGCGAGGAGCTGCGGGCCAGGCGGCC 240
DB 1778 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1719
QY 241 CTGGCCCGACATGGGGCCGCCACCTTCCCGAGAGCTCCCTGCGCCCTGAGCGGGAG 300
DB 1718 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1659
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 1658 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1599
QY 361 GCCTTCAGCATGACGGCTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 1598 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1539
QY 421 AAACAGCTGCACAGCTACTGCGCCATGTACACGCGAACCCAGGCGCTTGAGAGAGGCC 480
DB 1538 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1480
QY 481 CTGACAGCTGGCCACGCTGGGACACAGCCCGGGGAGCCCTTGGAGGTTGACCTGGAAGG 540
DB 1479 --GGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1422
QY 541 GCTGACATCCCTACGAGGACATCATAGCCATGAGATCTGAGGGGCTGCTTGGGAAGGC 600
DB 1421 GTGAGGCGGGGCTCGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
QY 601 GAGTCTGGGACCTGCGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
DB 1361 GGGGTAGAGGAGCTGAAGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1326

RESULT 5
US-08-910-647-1
Sequence 1, Application US/08910647
Patent No. 6251433
GENERAL INFORMATION:
APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
Polynucleotide Delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville

```
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/910,647
  FILING DATE:
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
    NAME: Fujita, Sharon M.
    REGISTRATION NUMBER: 38,459
    REFERENCE/DOCKET NUMBER: 1218.002
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (510) 923-2706
    TELEFAX: (510) 655-3542
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 9600 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

Query Match      17.9%; Score 114.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGTCTCAAGGCACACCTGAAGACACGCGG 60
DB 893 GAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 952
QY 61 GCGAGCTGAGGGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 953 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012
QY 121 GCGTGGAGCTGAGGCTGTGAGATGAGCTGAGCGCAAGAAAGAGGAGGAGGCTG 180
DB 1013 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072
QY 181 CTGCGGAGAGAGTGAACTGCTGAGCAGGAGGAGCTGAGGAGCTCGGGGCCAGG 240
DB 1073 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1132
QY 241 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCGGAGG 300
DB 1133 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1192
QY 301 CTGGAGCGGCTCGGGCCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 1193 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1252
QY 361 GCGTCCAGCATAGGGCTGCTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 1253 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1312
QY 421 AAACAGCTCAGCAGAGCTACCTGGCCATGTACACAGCGGAACAGCGCTGGAGA 480
DB 1313 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1371
QY 481 CTGCAAGCAGCTGGGAGCAGCGCCGGGAGGCCCTTGGAGGTTGACCTGGAAGG 540
DB 1372 --GGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1429
QY 541 GCTGACATCCCTACGAGACATCATAGCCATGAGATCTGAGGGGCTCCCTGGA 600
DB 1430 GTGGAGGGCGGGGTGAGGAGGAGTGTGGAGGGCGGGGTGAGGAGGAGTGTG 1489
QY 601 GAGTCTGGGACCTGSCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636
```

```
DB 1490 GGGGTAGGAGACGTGAAGAGCCAGGGGGGAAGTC 1525

RESULT 6
US-09-620-925-1
; Sequence 1, Application US/09620925
; Patent No. 6468986
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/620,925
; FILING DATE: 21-Jul-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/910,647
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1
```

```
Query Match      17.9%; Score 114.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGTCTCAAGGCACACCTGAAGACACGCGG 60
DB 893 GAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 952
QY 61 GCGAGCTGAGGGCTGAGCTGAGGAGCAGGAGCCAGGAGCTGAGGAGGAGGAGGAG 120
DB 953 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012
QY 121 GCGTGGAGCTGAGGCTGTGAGATGAGCTGAGCGCAAGAAAGAGGAGGAGGCTG 180
DB 1013 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072
QY 181 CTGCGGAGAGAGTGAACTGCTGAGCAGGAGCTGAGGAGCTCGGGGCCAGGCGCC 240
DB 1073 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1132
QY 241 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCGGAGG 300
DB 1133 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1192
```


REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 779
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-885-971-15

Query Match 17.9%; Score 114.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACGCCAAGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 2427 GAGGGCAGGACAGAGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGG 2486

QY 61 GCGAAGCTGGAGGCTGAGAGCTGAGGACCCAGGACCTGGAGGGGCGCCCTGCGCAACCAAG 120
DB 2487 CAGGAGCAGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2546

QY 121 GGCCTGGAGCTGGAGGCTGTGAGATGAGTGCAGCGCAAGAAAGAGGCGGAGCTG 180
DB 2547 GCGCAGGAGCAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2606

QY 181 CTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGAGGAGCTGCGGGCCAGCCGCC 240
DB 2607 CAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2666

QY 241 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGAGGCTGCGTGGGAAGGC 600
DB 2964 GTGAGGCGCGGGTCAAGAGAGTAGTGAGGCGCGGGTCAAGAGGAGTAGTGAGGCGCGC 3023

QY 601 GAGTCTGGGAGCTGGCACTGGGAGGCGAGGGCTCTC 636
DB 3024 GGGGTAGAGGACGTGAAGAGCCAGGGGGGAAGTC 3059

RESULT 9

US-08-087-783A-15
Sequence 15, Application US/08087783A
Patent No. 5547856
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,783A
FILING DATE: 13-Jul-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P0755779F1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-087-783A-15

Query Match 17.9%; Score 114.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACGCCAAGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 2427 GAGGGCAGGACAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGG 2486

QY 61 GCGAAGCTGGAGGCTGAGAGCTGAGGACCCAGGACCTGGAGGGGCGCCCTGCGCAACCAAG 120
DB 2487 CAGGAGCAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2546

QY 121 GGCCTGGAGCTGGAGGCTGTGAGATGAGTGCAGCGCAAGAAAGAGGCGGAGCTG 180
DB 2547 GCGCAGGAGCAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2606

QY 181 CTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGAGGAGCTGCGGGCCAGCCGCC 240
DB 2607 CAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2666

QY 241 CTGCGCCCGAGCATGTTGGGCGCCCGCCACTTCCCGAGGAGCTCCCTGCGCTGCGAGGGAG 300
DB 2667 CAGGAGCAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2726

QY 301 CTGGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCAAGCCATGACAGATGCTCTCG 360
DB 2727 CAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2786

QY 361 GCGTTCAGCATGAGCGGCTCTGTGGAAGGAGGAGAGAGGAAAGGTGATTGATTACCCAG 420
DB 2787 CAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAG 2846

QY 421 AAACAGCTGACAGAGAGTACGTGCCCATGTATACCGCGGAACCAAGCCCTGGAGAGGCC 480
DB 2847 CAGGAGCAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGA - 2905

QY 481 CTGACAGAGCTGGCAGCGGGGAGCAGCGCCGGGAGCCCTTGGAGGTTGACCTGGAAGGC 540

Db 2906 --GGAGGGGACAGGACAGGAGGGGACAGGAGGAGGTGGAGCGCGGGTTCGAGGAGCA 2963
QY 541 GCTGACATCCCTACGAGGACATACGACCTAGATCTGAGGGGCTGCTGGGAAGGC 600
Db 2964 GTGGAGGCGGGGTTCAGAGAGTAGTGGAGGCGGGGTTCGAGGAGGTAGTGGAGCGCGCC 3023
QY 601 GAGTCTGGGACCTGGCACTGGGAGGCGAGGGCTCTC 636
Db 3024 GGGGTAGAGGACGTGAAGAGCCAGGGGGGAAGTC 3059

RESULT 10
US-08-194-088B-15
; Sequence 15, Application US/08194088B
; Patent No. 5580963
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,088B
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallegos, R. Thomas
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: 755D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-2614
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-194-088B-15

Query Match 17.9%; Score 114.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACGCCAAGCTTAGCGAGATCTGGTCTCAAGGCACAGCTGAAGGACACGGCG 60
Db 2427 GAGGGGACAGGAGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGG 2486
QY 61 GGCAGAGTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120
Db 2487 CAGGAGCAGGAGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGG 2546
QY 121 GGCCTGGAGCTGGAGGTCTGTGAGAAATGAGTGCAGCGCAAGAAACGAGCGGAGCTG 180
Db 2547 GGGCAGGAGCAGGAGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGAG 2606
QY 181 CTCGGGGAAGAGGTGAACCTGCTGAGCAGGAGCTGCAAGAGCTCGCGGCCCGACCGCC 240

Db 2607 CAGGAGGGGACAGGACAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGGAGGGG 2666
QY 241 CTGGCCCCGACATGGGGCCCCCACCCTTCCCCGAGGACGTCCTTCCCTGCGCGGGAG 300
Db 2667 CAGGAGCAGGAGGGGACAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGGAG 2726
QY 301 CTGAGCGGGCTGCGGGGCGAGCTGCGGAGGAGCGGCAAGGCCATGACAGATGTCCTCG 360
Db 2727 CAGGAGGGGACAGGAGGGGACAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGG 2786
QY 361 GCGTTCAGCATGAGCGGCTGCTGGAGGAGGAGGAGGAGGAGGATTCAGTACCAG 420
Db 2787 CAGGAGGGGACAGGAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGGAGGGG 2846
QY 421 AAACAGCTGCAGCAGAGCTAGCTGGCCATGTACAGCGGAACCAAGCGCTGGAGAAGGCC 480
Db 2847 CAGGAGCAGGAGGGGACAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGGAG 2905
QY 481 CTGAGCAGCTGGGACGCTGGGGACAGCGCCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 540
Db 2906 --GGAGGGGACAGGAGGAGGGGACAGGAGGGGACAGGAGGGGACAGGAGGGGACAGG 2963
QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 600
Db 2964 GTGGAGGCGGGGTTCGAGGAGGTAGTGGAGGCGGGGTTCGAGGAGGGGACAGG 3023
QY 601 GAGTCTGGGACCTGGGCACTGGGAGGAGGGCTCTC 636
Db 3024 GGGGTAGGAGCGTGAAGAGCCAGGGGGGAAGTC 3059

RESULT 11
US-08-194-087-15
; Sequence 15, Application US/08194087
; Patent No. 5879910
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,087
; FILING DATE: 18-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregler, Girger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-194-087-15

Query Match 17.9%; Score 114.4; DB 2; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGCGG 60
Db 2427 GAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2486
QY 61 GGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2546
QY 121 GGCCTGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 2547 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606
QY 181 CTCGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db 2607 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2666
QY 241 CTGGCCCGGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2726
QY 301 CTGGAGCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 2727 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2786
QY 361 GGCCTTCCAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 2787 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2846
QY 421 AACAGCTTCACAGAGCTACGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2905
QY 481 CTCGAGCAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 2906 --GGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2963
QY 541 GCTGACATCCCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 2964 GTGAGGCGGGGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3023
QY 601 GAGTCTGGGAGCTGCGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
Db 3024 GGGGTAGAGGAGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3059

RESULT 12
PCT-US93-04648-15
; Sequence 15, Application PC/TUS9304648
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Natalie A., Mark, Melanie R
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04648
; FILING DATE: 19930517
; CLASSIFICATION:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755,779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US93-04648-15

Query Match 17.9%; Score 114.4; DB 5; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGCGG 60
Db 2427 GAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2486
QY 61 GGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2546
QY 121 GGCCTGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 2547 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606
QY 181 CTGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db 2607 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2666
QY 241 CTGGCCCGGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2726
QY 301 CTGGAGCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 2727 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2786
QY 361 GGCCTTCCAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 2787 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2846
QY 421 AACAGCTTCACAGAGCTACGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2905
QY 481 CTGCAGCAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 2906 --GGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2963
QY 541 GCTGACATCCCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 2964 GTGAGGCGGGGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3023
QY 601 GAGTCTGGGAGCTGCGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
Db 3024 GGGGTAGAGGAGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3059

RESULT 13
US-08-728-323A-1

Sequence 1, Application US/08728323A
Patent No. 5948676
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52269/JPW/MSK/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3489 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3489
US-08-728-323A-1

Query Match 13.9%; Score 88.8; DB 2; Length 3489;
Best Local Similarity 50.5%; Pred. No. 4.3e-09;
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCC 124
DB 1802 AGCAGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1861

QY 125 TGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGAGCCGCAAGAAACGAGGCGGAGCTGCTGC 184
DB 1862 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1921

QY 185 GGGAGAAGTGAACCTGCTGAGCAGCAGGAGCTGCGAGGAGCTGGGGCCCGCCCTGG 244
DB 1922 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1981

QY 245 CCGCGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCGAGCGGAGCTGG 304
DB 1982 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2041

QY 305 AGCGGCTCGGGCCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACACAGATGTCCTCGGGCT 364
DB 2042 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2101

QY 365 TCCAGCATGAGCGGCTCTGTGGAAGGAGGAGAGAGGAAGTATTGATTCACCAAGAAC 424
DB 2102 AGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2161

QY 425 AGCTGACAGCAGCTACGTGGCCATGTACACAGCGAACCCAGCGCTGAGAGAGCCCTGC 484
DB 2162 ATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2221

QY 485 AGCAGCTG 492
DB 2222 AGCAGCAG 2229

RESULT 14
US-09-298-568-1
Sequence 1, Application US/09298568
Patent No. 6322792
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballestas, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

Query Match 13.9%; Score 88.8; DB 4; Length 3489;
Best Local Similarity 50.5%; Pred. No. 4.3e-09;
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCC 124
DB 1802 AGCAGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1861

QY 125 TGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGAGCCGCAAGAAACGAGGCGGAGCTGCTGC 184
DB 1862 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1921

QY 185 GGGAGAAGTGAACCTGCTGAGCAGCAGGAGCTGCGAGGAGCTGGGGCCCGCCCTGG 244
DB 1922 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1981

QY 245 CCGCGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCGAGCGGAGCTGG 304
DB 1982 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2041

QY 305 AGCGGCTCGGGCCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACACAGATGTCCTCGGGCT 364
DB 2042 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2101

QY 365 TCCAGCATGAGCGGCTCTGTGGAAGGAGGAGAGGAAGTATTGATTCACCAAGAAC 424
DB 2102 AGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2161

QY 425 AGCTGACAGCAGCTACGTGGCCATGTACACAGCGAACCCAGCGCTGAGAGAGCCCTGC 484
DB 2162 ATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2221

QY 485 AGCAGCTG 492
DB 2222 AGCAGCAG 2229

RESULT 15
US-08-770-379-20/c

; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double.
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-770-379-20

Query Match 13.9%; Score 88.8; DB 2; Length 32207;
Best Local Similarity 50.5%; Pred. No. 5.7e-09;
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 65 AGCTGAGGCGCTGGAGCTGAGGACCCAGGACCTGAGGGGCGCCCTGCCGACCAAGGGCC 124
DB 20195 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20136
QY 125 TGGAGCTGGAGGTCTGTGAGATGAGCTCAGCGCAAGAAACGAGGCGGAGCTGTGTC 184
DB 20135 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20076
QY 185 GGGAGAGGTGACCTGCTGGAGCAGGAGCTCAGGAGCTCGGGCCCGCCCTGG 244
DB 20075 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20016
QY 245 CCGCGACATGGGGCGGCCACCTTCCCGAGGACGTCCTCCCTGCGAGCGGAGCTGG 304
DB 20015 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 19956
QY 305 AGCGGTGCGGCGCGAGCTGCGGAGGAGCGGCAAGCCATGACCAAGTGTCTCTGGGCT 364
DB 19955 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 19896
QY 365 TCCAGCATGCGGCTCGTGTGGAAGAGGAGGAGAGGTGATTCTAGTACCAAAAC 424
DB 19895 AGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 19836
QY 425 AGCTGAGCAGAGCTACGTGGCCATGTACCAGCGGAACAGGCGCTTGGAGAAAGGCCCTGC 484

Db 19835 ATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGC 19776
QY 485 AGCAGCTG 492
Db 19775 AGCAGCAG 19768
Search completed: June 15, 2003, 03:45:48
Job time : 39.4937 secs

Db 279 GAGCTGGTGCAAGGGCAGCAGCTGTGTGCTCTCGGGTGGCGCTGCGGAGGCGCGT 338
Qy 61 GGCAAGCTGGAGGGCCCTGAGCTGAGNACCAGAGACTTGGAGGGGCGCCCTCGGCACCAAG 120
Db 339 GCTACGCTGCGGGGTGAGTGGAGGGCGGTGCGGGGGTCTACAGGAGGCGCCGCGAGCTGG 398
Qy 121 GGCCTGGAGCTGGAGGTCTGTGAGAACTGAGCTGAGCGCAAGAAAGAGAGAGAGAGAGT 180
Db 399 GAGCTGGAGCTGGAGCTGTTCAGAGAGCTGTCCAGAGAGCTGAGAGACACCCGAGAGAGCTGAGCAG 458
Qy 181 CTGCGGGAGAGAGGTGAACCTCTCTGAGAGAGAGCTGACAGAGCTGCGGGC----- 230
Db 459 CTGCGGGAGAGAGCTGGGAGTGTGATGCTGAGGCGCGGAGCTCGGGAGGCCCCCTGTG 518
Qy 231 -----CCAGGCGCCCTGGCGCGGAGCATGGGGCGGCC 264
Db 519 CCACCTGCCACCGCTGACCCATTCTCTGCGCAGAGAGTGTAGAGGCCAAAGTGCAGCGG 578
Qy 265 ACCTTCCCGGAGGAGCTCCCTGCGCTGAGCGGGAGCTGGAGCGGCTGCGGGCCGAGCTG 324
Db 579 CGAGCAGCGGGTGGGGGAGCTTTCGGGCGCAGGTGGAGCGCATTCGGGTGGAGCTG 638
Qy 325 CGGAGGAGCGGCAAGGCCATGACAGATGCTCTGGGCTTCCAGCATGAGCGGCTCGTG 384
Db 639 CAGCGGAGCGGCGCGGGGTGAGGAGCAGCGGACAGCTTTGAGGGGAGCGCTGGCC 698
Qy 385 TGGAGGAGGAGAGGAGAGAGTGTTCAGTACAGAAACAGCTGCGAGCAGCTACGTG 444
Db 699 TGGCAGGAGAGAGGAGCAGTATCGCTTACCAGAGAGCAGCTGCGACAACTACATC 758
Qy 445 GCCATGTACAGCGGAACAGCGCTGAGAGAGCGCTGAGAGAGCTGGCAGCTGGGAGC 504
Db 759 CAGATGTACCGCGCAACCGAGCTAGAGAGAGCTGAGAGAGCTGAGCTGAGCTG 818
Qy 505 AGCGCGGGGAGCCCTTGGAGGTTGACCTG-----GAAGGGGCTGACATCCCTACGAGG 559
Db 819 GAGGCGGGGAGCTCGCTGACTGGGCTGCGGAGCAGCCCTGCTGCTGCTGGAGG 878
Qy 560 ACATCATAGCAGCTGAGATCTGAGGGGCTGCTGGAGGCGAGCTGCGGAGCTGGCAC 619
Db 879 AGATCATGCTACTGAGATCTAGGGCCCTGAGCAACAGCTCTGTAGGGAGCTCTGCCAG 938
Qy 620 TGGG 623
Db 939 AGG 942

RESULT 2
US-10-171-581-63
; Sequence 63, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Lineley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 63
; LENGTH: 5257
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB011124
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-63

Query Match 19.6%; Score 125.2; DB 9; Length 5257;
Best Local Similarity 70.9%; Pred. No. 3.8e-20;

Matches 166; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
Qy 286 GCCTTCAGCGGAGCTGGAGCGGCTGCGGCGGAGAGCGGAGGAGCGCAT 345
Db 3122 GCCTTCAGCGGAGGAGTGGGGCGGCTGCGAGCGAGCTGCGGGTGGAGCGGCGCGCG 3181
Qy 346 GACCAGATGTCCTCGGCTTCCAGCATGAGCGGCTCGTGTGGAAGAGGAGAGAGAGAG 405
Db 3182 GAGGCCAGGGTGCCAGCTTCGCCGAGGAGCGCGCTGTGGTGGAGAGAGAGAGAG 3241
Qy 406 GTGATTCAGTACCAGAAACAGCTGCGAGAGCTACGTGGCCATGTACAGCGGAACAG 465
Db 3242 GTGATCGAGTACCAGAAACAGCTGCGAGCTACGTGAGATGTACAGCGCAACAG 3301
Qy 466 CGCTTGAGAGAGGCGCTTCCAGCATGAGCGGCTGCGAGCTGGGAGAGCGGCGGAGCCC 519
Db 3302 CAGCTGAGCGCAGGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 3355

RESULT 3
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225

Query Match 12.4%; Score 79; DB 10; Length 2108;
Best Local Similarity 49.1%; Pred. No. 1.8e-09;
Matches 239; Conservative 0; Mismatches 245; Indels 3; Gaps 1;
Qy 44 AGCTGAAGGACACGCGGGGCAAGCTGGAGGCGCTTGGAGCTGAGAGCCAGGAGCTGGAGG 103
Db 898 AGGGGAGCTGAAGCACCTGGATCAGCAGGAGAGAGCAGCAGAGCTCCAGAGCAGCAGA 957
Qy 104 GCGCCCTGGCACCAGGGCTTGGAGCTTGGAGTCTGTGAGATGAGCTGCGAGCGCAGA 163
Db 958 TGGGGCAGCTGAAGCACCTTGGAGCAGGAGGAGGGGCGAGCCCTAAGCATCTGAGCAGCAGG 1017
Qy 164 AGACAGAGGCGGAGCTCTCGGGAGAGAGTGAACCTGCTGGAGCAGGAGCTGCGAGGAGC 223
Db 1018 AGGGGCACTGGAGCAGCTGAGGAGCAGGAGGAGGAGCTGAACACCTTGGAGCAGCAGG 1077
Qy 224 TCGGGGCGCCAGGGCGCCCTGGCGCCGACATGGGGCGCGCCACCTTCCCGAGAGCGTCC 283
Db 1078 AGGGGAGCTGGAGCACCTTGGAGCAGCAGGAGGAGGAGCTGGGGCTCCAGCAGCAGG 1137
Qy 284 CTGCCCTGCGAGGAGCTGGAGCGGCTTGGGGCGGAGCTGCGGGAGGAGCGGCGCAAGGCC 343
Db 1138 TGCTGCAGCTGAAGCAGCTAGAGAGCAGCAGGAGGAGCGCCAAAGCACTTGGAGAGGAGG 1197
Qy 344 ATGACAGAGTCTCTCGGGCTTCCAGCATGAGCGGCTGTGTGGAAGAGGAGAGAGAGA 403
Db 1198 AGGGGAGCTGAAGCACCTTGGTGCAGCAGGAGGAGGAGCTGAACATCTGTGAGCAGG 1257
Qy 404 AGGTGATTCAGTACCAAGAAACAGCTGCGAGCAGAGCTACGTGGCCATGTACAGCGGAACC 463
Db 1258 AGGGGAGCTGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1317

Qy	464	---	AGCCCTGGAGAGGCCCTTCGACGAGCTGGCAGCOTGGGGACAGCGCCGGGGAGCCCT	520
Db	1318	TGAAGCACCTAGAGGACAGGAGGGACAGCTGAAGCATCTGGAGCAGCAGCAGGGGCAGT	1377	
Qy	521	TGGAGGT	527	
Db	1378	TGGAGGT	1384	

RESULT 4
 US-09-969-852-4
 ; Sequence 4, Application US/09969852
 ; Patent No. US20020137211A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Tianyan
 ; APPLICANT: Liu, Huifen
 ; APPLICANT: Li, Wei
 ; APPLICANT: Zhao, Libin
 ; TITLE OF INVENTION: A METHOD FOR ESTABLISHING AN EXPRESSION SYSTEM OF SPIDER DRAGLINE
 ; TITLE OF INVENTION: GENE IN BOMBYX MORI
 ; FILE REFERENCE: LIU=65
 ; CURRENT APPLICATION NUMBER: US/09/969,852
 ; CURRENT FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: CN01106406.4
 ; PRIOR FILING DATE: 2001-01-02
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1852
 ; TYPE: DNA
 ; ORGANISM: Nephila clavipes
 US-09-969-852-4

Query Match	12.2%	Score 77.8	DB 10	Length 1852
Best Local Similarity	47.3%	Pred. No. 3.3e-09		
Matches 235	Conservative 0	Mismatches 262	Indels 0	Gaps 0
Qy	51	GGACACGGGGCAAGCTGAGGCGCTGGAGGCTGAGGACCCAGGACCTGAGGGCGCCCT	110	
Db	1045	GGACAAAGTGCAGCAGCAGCAGCAGCTGGAGGTGCTGGACAAGAGGATATGAGAGTCTT	1104	
Qy	111	GGCACCACAAAGGCTGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAGAACGA	170	
Db	1105	GGAAAGCCAAAGTGCAGACACAGGTGGATTAGTTAGTGGACAAGGTGCAGTCAGCAGACGA	1164	
Qy	171	GGCGGAGCTCTGCGGAGAAAGTGAACCTGCTGGAGCAGAGCTGCAGAGCTGCGGGC	230	
Db	1165	GCAGCTGAGGTGCTGGACAAGGAGGATACGGTGGTCTTGGTGGACAAGGTGCGGACAA	1224	
Qy	231	CGAGGCGCCCTGGCCCGCACATGGGGCGGCCACCTTCCCGCAGGACGTCCTGCGCCT	290	
Db	1225	GGAGGCTATGGAGGACTTGGAAAGCCAAAGTTCTGGTGCAGGAGGATTAGGTGGACAAGGT	1284	
Qy	291	GCAGCGGAGCTGAGCGGCTGCGGGCCGAGCTGCGGAGAGCGGCAAGCCATGACCA	350	
Db	1285	GCAGGTGCACGACGACGACGAGCTGGAGGTGCTGGACAAGAGGATTAGGTGACAAAGT	1344	
Qy	351	GATGTCTCTCGGGTTCACGATGAGCGGCTCGTGTGGAAAGGAGGAGAGAGGTGAT	410	
Db	1345	GCTGGACAAGGAGCTGGAGCAGCCGCTGCAGCAGCTGGTGGTTAGACAAGGAGGATAT	1404	
Qy	411	TCAGTACCAAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCGCGGAACAGCGCCT	470	
Db	1405	GGAGGTCTTGGAAAGCCAAAGTGTGGACGAGGTGGACAAGGTGCAGCGCAGCGCAGCA	1464	
Qy	471	GGAGAGGCGCTGCAGCAGCTGGCACCTGGGGACAGCGCCGGAGGCCCTTGGAGGTTGA	530	
Db	1465	GCAGCGGAGGTGCTGGACAAGGAGGATATGGTGGTCTTGGTGGACNAGGTGTGGCCGA	1524	
Qy	531	CCTGGAGGGGCTGACA	547	
Db	1525	GCTGGATTAGGTGGACA	1541	

RESULT 5

```

US-009-854-133-337
; Sequence 337, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND
; TITLE OF INVENTION: THE THERAPY AND
; FILE REFERENCES: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 337
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(422)
; OTHER INFORMATION: n = A,T,C or G
US-009-854-133-337

```

Query Match 12.1%; Score 77.4; DB 9; Length 422;
Best Local Similarity 49.6%; Pred. NO. 3.8e-09;
Matches 198; Conservative 0; Mismatches 201; Indels

143	Qy	AGAATGAGCTGCAGCCGACAAGAGAACAGAGCGGAGCTGTCTGGGAGAGAGGTGAACCTTGC	202
6	Db	AGGAACAGTTTACAGCAGCAGCAGCAACAGCAGAGCTGTTTGCACAACAGCAGCAGGAACAATTGC	65
203	Qy	TGCAGCAGGAGCTGCAGGAGCTCGGGCCAGAGCCGCTTGGCCCGGACACATGGGGCCG	262
66	Db	AGCAGACAACTTGCAGCCTCTCCCTTGAGCCCGAGAGAGGAAAGGTGTGAGCTTGG	125
263	Qy	CCACCTTCCCGGAGGACGTCTCCCTGCAGCGGGAGCTGGAGCGGTGCGGGCCGAGC	322
126	Db	AGCTCATCGGTGACCTGGGTCTCAGACAGAGAGCTTGGAGCAGCGGCAGAGTTGG	185
323	Qy	TGCGGAGGAGCGCAAGGCATAGCAAGATGTCTCTCGGGCTTCCAGCATAGCGGTCTG	382
186	Db	AGCGGCAGCAGGAGCTTGAACGGCAGCAGGAGCAGCGGACGTGCAGCTCAAATGTCAGG	245
383	Qy	TGTTGGAAAGGAGAGAGGAGAGGTGATTTCAGTACCAGAAACACCTGCAGCAGAGCTACG	442
246	Db	AGGAGCTGCAGCAGCTGGAGCAACAGCTGGAGCAGCAGCAGCAGCTGGAGCAGCAGG	305
443	Qy	TGGCCATGTACACAGCGGAACACAGCGCTTGGAAAGGCCCTTGCAGCAGCTGGCACGTGGGG	502
306	Db	AGTGCAGCTGGAGCTGACCCCGTGGAGTAGGGCGCCACAGCAGCAGGAGGTGAGCTGG	365
503	Qy	ACACGCGCGGGGAGCCCTTGGAGGTTGACCTGGAAAGGG	541
366	Db	AGCTGACCCCGGTGCAGCCGGAGCTGCGAGCTTGGAACTGG	404

DEC 11 1954

RESOLI 6
US-09-738-973-337
; Sequence 337, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Alqaate, Paul A.

APPLICANT: Secrist, Heather
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Benson, Darin R.
APPLICANT: Elliot, Mark
APPLICANT: Mannion, Jane
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.475C9
CURRENT APPLICATION NUMBER: US/09/738,973
CURRENT FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 587
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 337
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapien
NAME/KEY: misc feature
LOCATION: (1)-(422)
OTHER INFORMATION: n = A,T,C or G
US-09-738-973-337

Query Match 12.1%; Score 77.4; DB 10; Length 422;
Best Local Similarity 49.6%; Pred. No. 3.8e-09;
Matches 198; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
QY 143 AGATGAGCTGAGCGGCAAGAAACAGAGCGGAGCTCTCTGGGAGAGGTGAACCTGC 202
Db 6 AGGAACAGTTACAGCAGCAGCAGCAACAGCAGCTTTGCAACAGCAGCAGCAAAATTGC 65
QY 203 TGGAGCAGAGCTGAGGAGCTGGGGCCAGGCGCCCTGGCCCGCACATGGGGCGC 262
Db 66 AGCAGCAACACTGAGCAGCTCTCCCTGGAGCCCGAGGAGGAGAGGTGAGCTGG 125
QY 263 CCACCTTCCCGAGAGCTCCCTGCTGAGCGGAGCTGGAGCGGCTGGGGCCGAGC 322
Db 126 AGCTCATCGCGGTGACCTTGGGTGAGAGCAGGAGCTGGAGCAGCGGGCAGAGTTGG 185
QY 323 TGGCGGAGAGCGGCAAGCCATGACAGATGTCTCTGGGTTCAGCATGAGCGGCTCG 382
Db 186 AGCGCAGCAGAGCTGGAACCGCAGCAGGAGCAGCGCAGCTGAGCTCAAACTGCAGG 245
QY 383 TGTGAAGCAGGAGAGCAGAGGTGATTGATCAGTACAGAAACAGCTGAGCAGAGCTACG 442
Db 246 AGGAGCTGAGCAGCTGGAGCAACAGCTGGAGCAGCAGCAGCAGCTGGAGCAGCAGG 305
QY 443 TGGCCATGTACCGCGGAACAGCGCCCTGGAGAGCGCCCTGAGCAGCTGCGACGCTGGGG 502
Db 306 AGGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGG 365
QY 503 ACAGCGCGGGAGCGCTTGGAGTTGACCTGGAGAGGGG 541
Db 366 AGCTGACCCCGCTGAGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 404

RESULT 7

US-09-764-891-6987/c
Sequence 6987, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6987
LENGTH: 22595
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (9701)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-6987

Query Match 12.1%; Score 77.2; DB 9; Length 22585;

Best Local Similarity 46.9%; Pred. No. 5.2e-09;

Matches 241; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 13 AAGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACACGCGGGGCAAGCTGAG 72
Db 10356 AAGGAGATGAGATGACAGGAGGAGGACGACAGATGACAGGAGGAGCGCAGGTCAGAGGGG 10297
QY 73 GGCTGAGCTGAGGACCCAGGACCTGAGGAGGCGCCCTGCGCACCAGAGGGCTGAGGTG 132
Db 10296 ACGCAGATGACAGGAGGAGGAGCGGCTGACAGGAGGAGATGACAGGAGGAGTACAG 10237
QY 133 GAGTCTGTGAGAAATGAGCTGCAGCGCAAGAAACAGCGGAGCTGCTCGGGGAGAA 192
Db 10236 GAGGGTTGGGCTGACAGGAGGAGGACACAGATGACAGGAGGAGGAGGAGTGCAGAGGCG 10177
QY 193 GTGAACCTGCTGAGCAGGAGCTGACAGAGCTGCGGCCCGAGCGCCCTGGCCCGGAC 252
Db 10176 ACGGGCTGCAGAAAGAGATGACAGATGACAGGAGGAGTGCAGATGACAGGAGGAGTGCAG 10117
QY 253 ATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCTGCGGAGCTGAGGAGGAGCTGAGCGGCTG 312
Db 10116 ATGACAGAGGGGATGACAGGTGACAGGAGGAGGAGCGGGCTGACAGGAGGAGGAGGAGTGCAG 10057
QY 313 CGGGCCGAGCTGCGGGAGGAGCGCAAGGCGCATGACAGATGCTCTCGGGCTTCCAGCAT 372
Db 10056 GGNACGGGCTACAGGAGGAGATGACAGATGACAGGAGGAGGAGTGCAGATGACAGGAGTGTGCG 9997
QY 373 GAGCGCTCTGTGGAAGGAGGAGAGAGAGTATTGATACCAACAAACAGCTGAG 432
Db 9996 GGGCTGACAGGAGGAGCGCAGATGACAGGAGGAGCGTGGGGCTGACAGGAGGAGGAGCAGATC 9937
QY 433 CAGAGCTACGTGGCCATGTACAGCGGAACAGCGCTGAGAGAGGCGCTGAGAGAGGCGCTGAGCAGCTG 492
Db 9936 CAGAGGAGGAGCGGGTGCAGGAGGAGGAGCGGGTGCAGGAGGAGGAGGAGGAGGAGGAGGAG 9877
QY 493 GCACGTGGGACAGCGCGGGAGCGCTTGGAGG 526
Db 9876 GGNACGGGCTGACAGGAGGAGGAGCAGAGTGCAGG 9843

RESULT 8

US-09-954-456-1601
Sequence 1601, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720

US-09-738-973-157

Query Match 11.5%; Score 73.8; DB 10; Length 2313;
Best Local Similarity 50.6%; Pred. No. 2.8e-08;
Matches 208; Conservative 0; Mismatches 197; Indels 6; Gaps 1;

QY 77 TGGAGCTCAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCTGGAGCTGGAGG 136
DB 960 TGGCGATCCGCTGACGAGCCGAGGGGAAGAGCTCCCTACACAGAGACTGAAGCAGAAAC 1019

QY 137 TCTGTGAGATGAGCTGCAGGCGCAAGAAACGAGGCGGAGCTGCTCGGAGAGAGTGA 196
DB 1020 GCGCGAGCAGCGGAGCAGCGGAGCGCGCGCGG-----CCAAGGAAGAGGAGC 1073

QY 197 ACCTGCTGAGCAGGAGCTGCAGGAGCTGCGGGCGCCAGCGCCCTGCGCCCGCACATGG 256
DB 1074 TGCTGCGGCTGCAGCAGCTGCAGGAGGAGAGGAGCGGAAGCTGCAGGAGCTGGAGCTGC 1133

QY 257 GCGCGCCACCTTCCCGCAGGACGTCCTGCGCCCTGCGAGCGGAGCTGGAGCGGCTGCGGG 316
DB 1134 TGCAGGAGCGCAGCGGCGGAGCGGCTGTCTGCGAGGAGGAGGAACGGCGCCGCA 1193

QY 317 CCGAGCTCGGGAGGAGCGCAAGSCCATGACAGATGCTCTCGGGCTTCAGAGCATGAGC 376
DB 1194 GCCAGCAGCGAGCTGCAGCAGCGCTCGAGGGCCAACTCGCGAGSCGAGCAGGCCC 1253

QY 377 GGCTGCTGGAAGAGGAGAGAGAGAGGTGATTTCAGTACAGAAACAGCTGCAGCAGA 436
DB 1254 GGGCTCCATCGAGCTGAGATGGAGCTGGAAGGAGGAGGAGGCTGCCCGCAGCGGCGAGC 1313

QY 437 GCTAGCTGCGCATGTACCAGCGGACCAAGCGCTTGGAGAAGCGCTCGAGC 487
DB 1314 GCATCAAGAGCTGGAGGAGATGACAGCGGTTGCAGGAGCGCTTGCAC 1364

RESULT 11

US-09-764-868-12
; Sequence 12; Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-12

Query Match 11.5%; Score 73.8; DB 9; Length 2314;
Best Local Similarity 50.6%; Pred. No. 2.8e-08;
Matches 208; Conservative 0; Mismatches 197; Indels 6; Gaps 1;

QY 77 TGGAGCTCAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCTGGAGCTGGAGG 136
DB 968 TGGCGATCCGCTGACGCGCGGAGGAGAGAGCTCCCTTACACAGAGCTGAAGCAGAAAC 1027

QY 137 TCTGTGAGATGAGCTGCAGGCGCAAGAGACGAGGCGGAGCTGCTCGGAGAGAGTGA 196
DB 1028 GCGCGAGCAGCGGAGCAGCGGAGCGCGGCGCGCGCGG-----CCAAGGAAGAGGAGC 1081

QY 197 ACCTGCTGAGCAGGAGCTGCAGGAGCTGCGGGCGCCAGCGCCCTGCGCCCGCACATGG 256
DB 1082 TGCTGCGGCTGCAGCAGCTGCAGGAGGAGAGGAGCGGAAGCTGCAGGAGCTGGAGCTGC 1141

QY 257 GCGCGCCACCTTCCCGCAGGACGCTCTGCGCCCTGCGAGCGGAGCTGCAGCGGCTGCGGG 316
DB 1142 TGCAGGAGCGCAGCGGCGGAGCGCGGAGCGGCTGCTGCAGGAGGAGGAGGAACGGCGCCGCA 1201

QY 317 CCGAGCTCGGAGGAGCGGCAAGGCCATGACAGATGTCTCTCGGGCTTCAGCATGAGC 376
DB 1202 GCCAGACCCGAGAGCTGCAGCAGCGCTCGAGGGCCAACTGCGGAGCGGAGCAGGCC 1261

QY 377 GGCTGCTGTGCAAGGAGGAGGAAGAGAGTGTTCAGTACCAAGAAACAGCTGCAGCAGA 436
DB 1262 GGGCTTCCATGCAGGCTGAGATGGAGCTGAAGGAGGAGGAGTGTCCCGCAGCGGAGC 1321

QY 437 GCTACGTGGCCATGTACCAGCGGAACCGAGCGCTTGGAGAAGGCCCTTCGAGC 487
DB 1322 GCATCAAGGAGCTGGAGGAGATGCAGCAGCGGTTGCAGAGGAGGCCCTCAAC 1372

RESULT 12

US-09-925-299-215
; Sequence 215; Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 215
; LENGTH: 3323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-299-215

Query Match 10.9%; Score 69.6; DB 9; Length 3323;
Best Local Similarity 48.1%; Pred. No. 2.7e-07;
Matches 232; Conservative 0; Mismatches 244; Indels 6; Gaps 1;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTTGGAGGCGCCCTCGCACCAAGGGCC 124
DB 1198 AGCTGGTGGCGCGGAGCAGGAGATCACGGCTGTGTCAGGCAAGCATGCAGGCCAGCTACC 1257

QY 125 TGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGACGAGCGGAGCTGCTGC 184
DB 1258 GGGAGCACGTGAAGAGGTCAGCAGCTGCAGGGCAAGATCCCGAGCTTTCAGGAGCAGC 1317

QY 185 GGGAGAAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCGCCAGCGCCCTTGG 244
DB 1318 TGGAGATGGCCCCAACACAGCAGCTGCGCGCTGCAGCAGG-----AGNACTCCATCC 1371

QY 245 CCGCGCATATGGGCGCCGCCACCTTCCCGAGGACGTCCTCTGCGCTCGAGCGGAGCTGG 304
DB 1372 TCGGGGATGCTTGAACACAGGCCACAGAGCCAGGCTGGAGAGCAAGCAGAAACGAGAGTGG 1431

QY 305 AGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCAAGGCCATGACCAATGCTCTCGGGCT 364
DB 1432 CCAAGCTTCGGCAGGAGCTCAGCAAGGTTCAGCAAGAGCTGGTGGAGAAGTCAGAGGCTG 1491

QY 365 TCACAGCATGACGCGCTCGTGTGGAAGGAGGAGGAGGAAGGTGATTTCAGTACCAAGAAAC 424
DB 1492 TCGGCGAAGATGACAGCAGCGGGAAGCTCTGGAGCAAGGAGCTGCTTCAGAAAGC 1551

QY 425 AGCTGCAGCAGAGCTACCTGGCCCATGTACCAGCGGAACCAAGCGCTTGGAGAGGCCCTTGC 484
DB 1552 AGGTCTTCAGCTCAGCGCGTCCCAAGGAGAGTGAAGGAGGCCCTTCAGAAAGCGCCCTGG 1611

QY 485 AGCAGCTGGCAGCTGGGGAGCAGCGCCGGGAGCCCTTGGAGGTTGACTTGAAGGGGCTG 544
DB 1612 ACAGAGTCAGCCGGGAGCTGTGCCACACGACGACGACGCCAGCTCCGGGCGGATG 1671

QY 545 AC 546

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 889.672 Seconds
(without alignments)
11632.279 Million cell updates/sec

Title: US-09-513-888c-1_COPY_4912_5550

Perfect score: 639

Sequence: 1 gaggtgaacccaaggtag.....tggaggcagggtctctcccg 639

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	175.8	27.5	247	9	AL135213
2	171.2	26.8	326	10	AW457817
3	159.4	24.9	795	13	BI335611
4	131.8	20.6	562	12	BF079806
5	128.8	20.2	824	13	BI655319
6	125.2	19.6	1053	14	BQ431916

7	121.2	19.0	918	12	BG385903
8	120.4	18.8	482	13	BM152503
9	120	18.8	453	14	BM855056
10	120	18.8	901	12	BF317160
11	120	18.8	926	12	BG035858
12	120	18.8	957	12	BG170653
13	118.6	18.6	923	12	BG750395
14	118.4	18.5	469	13	BM193264
15	118.2	18.5	450	14	BQ694854
16	117.2	18.3	705	13	BI104307
17	117	18.3	357	14	BM854246
18	116.2	18.2	806	13	BM005722
19	115.2	18.0	914	13	BI766138
20	114.8	18.0	977	12	BF032717
21	114.4	17.9	799	13	BI146842
22	114.2	17.9	898	14	BQ646816
23	114.2	17.9	940	12	BG830547
24	113	17.7	718	17	AZ972907
25	112.4	17.6	704	13	BM042462
26	112.2	17.6	342	13	BM106435
27	111.4	17.4	883	14	BQ946535
28	110.8	17.3	858	13	BI250802
29	110.4	17.3	287	14	BM818585
30	107.4	16.8	839	13	BI752956
31	107	16.7	811	17	CNS038X0
32	106.6	16.7	973	9	AL551801
33	106.4	16.7	432	12	BF313433
34	106.4	16.7	1106	14	BQ648477
35	105.6	16.5	587	12	BF309857
36	105.2	16.5	650	17	AZ337339
37	102.4	16.0	705	17	AZ640388
38	101	15.8	751	10	BE312985
39	101	15.8	932	14	BQ879505
40	100.2	15.7	395	13	BM146503
41	100	15.6	564	13	BJ098506
42	94.8	14.8	486	9	AI605748
43	94.8	14.8	811	12	BG866479
44	94.2	14.7	713	17	AZ346186
45	93.8	14.7	535	10	AW420993

ALIGNMENTS

RESULT 1

AL135213

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL135213 247 bp mRNA linear EST 25-FEB-2000
DKFZp762E2313.r1 762 (synonym: hmel2) Homo sapiens cDNA clone
DKFZp762E2313.5', mRNA sequence.

AL135213

AL135213.1 GI:6603400

EST

human.

Homo sapiens

1 (bases 1 to 247)

(bases 1 to 247)

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No s1 sequence available.

This clone (DKFZp762E2313) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No s1 sequence available.

This clone (DKFZp762E2313) is available at the RZPD in Berlin.

Meleagris.

REFERENCE 1 (bases 1 to 795)
 AUTHORS Harry D.E., Zaitlin, D., Marini, P.J. and Reed, K.M.
 TITLE A first-generation map of the turkey genome
 JOURNAL Unpublished (2001)
 COMMENT Contact: Reed KM

Department of Veterinary Pathobiology,
 University of Minnesota
 295 AS/VN, 1988 Fitch Ave, St. Paul, MN 55108, USA
 Tel: 612 624 1287
 Fax: 612 625 0204
 Email: reedx054@cc.umn.edu
 Seq primer: TCGAAGTTCCTCCCTCACAATAAGG
 POLYA-No.

FEATURES

Location/Qualifiers

1..795
 /organism="Meleagris gallopavo"
 /db_xref="taxon:9103"
 /clone="Nte083"
 /clone_lib="Turkey Lambda ZAP Library"
 /sex="unknown"
 /tissue_type="whole embryo"
 /dev_stage="24-day"
 /note="Vector: Lambda ZAP Library ZAP; RNA was isolated from a 24-day turkey embryo using a commercially available kit (Stratagene). cDNA library was constructed using a Lambda ZAP (SK-, Stratagene)."

BASE COUNT 188 a 232 c 214 g 161 t

Query Match 24.9%; Score 159.4; DB 13; Length 795;

Best Local Similarity 77.5%; Pred. NO. 5.7e-19;

Matches 193; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 334 CGGAAGCCATGACCATGCTCCGGCTTCAGCATGAGCGCTCGTGTGAAGAG 393

Db 1 CGGACACGACGAGCATGAGTCCAGCTTCAGCAGGAGCGGAGCATGAGGAG 60

Qy 394 GAGAAGGAGAGGTGATTTCAGTACAGAAACAGCTGCAGCAGAGCTAGCTGGCCATGTAC 453

Db 61 GAAAAGAGAGGTGATCCATACAGAAAGCAGTTGCAGCAGAGCTACCTGCATGTAC 120

Qy 454 CAGCGAACCAGCGCTCGGAAGGCCCTGCAGCAGTGGCAGCTGGGAGCAGCGCCGGG 513

Db 121 AAGAGAACAGAACCTGGAGAAGATGCTGCAGCAGCTCGCGGGGGGAGGTGCAAG 180

Qy 514 GAGCCCTGGAGTTGACTTGAAGGGGCTGCATCCCTACGAGGACATCATAGCCACT 573

Db 181 GAGCCCAFTGAGCTGGACATCCCGCGCTGACGTGCTCCCTACGAGGACATCATAGCCACT 240

Qy 574 GAGATCTCA 582

Db 241 GAGATCTCA 249

RESULT 4

LOCUS BF079806 562 bp mRNA linear EST 18-OCT-2000
 DEFINITION 230497 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF079806
 VERSION BF079806.1 GI:10873636
 KEYWORDS EST.

SOURCE

Sus scrofa

pig.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 562)

Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
 and Keele, J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

TITLE

Unpublished (2000)

JOURNAL

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCCAGTCACGACG

Plate: 47 row: K column: 8

Seq primer: ATTAGGTGACACTATAG.

FEATURES

Location/Qualifiers

1..562
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 2P1G"
 /tissue_type="pooled"
 /lab_host="DH108"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from testis, ovary, and

endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 106 a 165 c 226 g 65 t

ORIGIN

Query Match 20.6%; Score 131.8; DB 12; Length 562;

Best Local Similarity 54.9%; Pred. No. 4.5e-14; Indels 36; Gaps 1;

Matches 308; Conservative 0; Mismatches 217;

Qy 20 GCGAGATCTCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAAGCTGGAGGCGCTGG 79

Db 1 GCGAGCTGGTAGCCTCGAGTGGCGCTACGGGAGGCGCGGCGAGCGCTGCGGCTCAGTG 60

Qy 80 AGCTGAGACCCAGGACCTGGAGGCGCCTTGCGCACCAAGGGCTGTGAGCTGAGGTCT 139

Db 61 AGGCGCGGGCGGGGCTCCAGGAGCGCGCCGCAACACGCGAGCTGTGAGCTGGAGGCGCT 120

Qy 140 GTGAGATGAGTCTCAGCGCAAGAAAGACGAGGCGGAGCTGCTCGGGGAGAGTGAACC 199

Db 121 GTTCCCAAGAGCTGCACGGCATGCCAGAGGCTGAGCGCTCGGAGAGAAAGCGGGAC 180

Qy 200 TGCTGGAGCAGGAGCTGCAGGAGCTGCGGGG- 230

Db 181 AGTTGGACACCGAGGCGGTGACTCCGGGAACCCCTTGTGCCACCTGCCACTGCCGACC 240

Qy 231 -----CGAGGCGCCCTGGCGCGCAGCATGGGGCGCCACCTTCCCGAGGACGTCC 283

Db 241 CATTCCTCTGGCAGAGAGCGATGAAGCAAGGCACAGCGGGCAGCTGCTGGGTTGGGG 300

Qy 284 CTGCGCTGCGAGGAGCTGGAGCGGCTGCGGGCGGAGCTGCGGAGAGCGGCAAGGCC 343

Db 301 CGAGCCTGCGGGGCCAGGTGGAGCGCTCTCGGGGAGATCTGACGCGGAGCGGGCGGTG 360

Qy 344 ATGACACAGATGTCCTCGGGCTTCAGCATGAGCGGCTCGTGTGGAAGAGAGAGAGA 403

Db 361 GCGAGGAGCAGCGGACAGCTTTAGGGGAGCGGCTGGCTGCGAGCGCGGAGAGGAGC 420

Qy 404 AGGTGATTTCAGTACCAAGAACAGCTGCAGCAGAGCTACGTGGCCATGTACAGCGGAGACC 463

Db 421 AGGTGATCCGCTACCAAGACAGCTGCAGCAGCACTACATCCAGATGTACCGAGCGCAACA 480

Qy 464 AGCGCTTGGAGAGGCGCTTCAGCAGCTGGGAGACAGCGCGGGAGCGCCCTTGG 523

Db 481 GGCAGCTGGAGCAGGAGCTACAGCAGCTCAGCTGGAGCTGGAGGCGCGGAACTCGCCG 540

Qy 524 AGGTGACCTGGAGAGGGCTG 544

Db 541 ATCTGGCCTGGCGAGCCAG 561

RESULT 5

```

BI655319          824 bp      mRNA      linear      EST 12-SEP-2001
LOCUS             603283348F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5327843 5',
DEFINITION        mRNA sequence.
ACCESSION         BI655319
VERSION           BI655319.1 GI:15569555
KEYWORDS          EST.
SOURCE            house mouse.
ORGANISM          Mus musculus.
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS           1 (bases 1 to 824)
TITLE            NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL           National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
                  Ph.D.
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone Distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM11831 row: o column: 12
                  High quality sequence stop: 823.
                  Location/Qualifiers
                    1..824
                     /organism="Mus musculus"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5327843"
                     /clone_lib="NCI_CGAP_Mam4"
                     /tissue type="tumor, gross tissue"
                     /dev stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and Differentiation 7, 3-11 (1996)."
```

```

BASE COUNT       156 a 251 c 295 g 122 t
ORIGIN
Query Match      20.2%; Score 128.8; DB 13; Length 824;
Best Local Similarity 55.3%; Pred. No. 1.6e-13;
Matches 321; Conservative 0; Mismatches 217; Indels 42; Gaps 2;
QY 58 CCGGCAAGCTGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACC 117
DB 42 CCGGCCACCTTGGCGGTCAGTGAGGGCGGTGCCAGGGCCCTGCAGAGGGCAGCCGAGCT 101
QY 118 AAGGCGCTGGAGCTGAGGTCTGTGAGATGAGCTGCAGCGCAAGAACGAGCGGAG 177
DB 102 CCGGAGCAGGAGCTGGAGGCTCTGCGAGGAGCTTCAGCGGTACCGCCAGGAGGCTGAG 161
QY 178 CTGCTCGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGC----- 230
DB 162 CGGCTCGGGAGAGGCGGGGCAATTGGATGCCGAAGCATCGGACTCGGGATCCCT 221
QY 231 -----CGAGCGCCCTGGCCCGCCGACATGGGGCG 261
DB 222 GTACCCCTGCCACACCGACCCGTTCTCTTGGCAGAGAGTGTAGGCCCAAGGTGCAG 281
QY 262 CCACCTTCCCGAGAGACTCCCTGCCTGTCAGCGGGAGCTGCAGCGCTGCGGGCCGAG 321
DB 282 CCGGAGAGCGCTGGGCGAGGGGGTAGCTGCGGGCTCAGGTGGAACGCTGCGCCAGAA 341
QY 322 CTGCGGGAGGAGCGGCAAGCCATGACACAGATGTCTCTGGGCTTCCAGATGAGCGGCTC 381
DB 342 CTGCAGCGGAGCAGCGGCGAGGGGACGAGCAACGGGACAGCTTTGAGGGGAGCGGCTG 401
```

```

QY 382 GTCTGGAGGAGGAGGAGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGCTAC 441
DB 402 GCCTGGGAGGAGGAGGAGGAGGAGGTGATCCGCTACCAGAGGAGCTGCAGCAACTAT 461
QY 442 GTGGCCATGTACCAAGCGGAAACCAAGCGCTGGAGAAAGCCCTGCAGCAGCTGGCAGTGGG 501
DB 462 ATCCAGATGTACCGGCTTACCGGAGCTGGAGCAGGAACTGCAGCAACTGAGCCTGGAG 521
QY 502 GACAGCCCGGGAGAGCCCTTGGAGTTGACTGGAAGGG-----GCTGACATCCCCCTAC 555
DB 522 CTGGAGCCCGCGAGCTTGGGATTTGGGCTGGCGGAGTCGGCCCTTCATCTGTCTG 581
QY 556 GAGGACATCATAGCCACTGAGATCTGAGGGGCTGCGCTGGG 595
DB 582 GAGGAGATCACTGCCACTGAGATCTAGGGCTCTGCAAGGG 621

RESULT 6
LOCUS             B0431916
DEFINITION        AGENCOURT 7998255 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154088
                  5', mRNA sequence.
ACCESSION         B0431916
VERSION           B0431916.1 GI:21170992
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS           1 (bases 1 to 1053)
TITLE            NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL           National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue Procurement: ATCC
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  Clone Distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM13494 row: j column: 09
                  High quality sequence stop: 378.
                  Location/Qualifiers
                    1..1053
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6154088"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Library constructed by Life
                     Technologies."
BASE COUNT       179 a 349 c 377 g 148 t
ORIGIN
Query Match      19.6%; Score 125.2; DB 14; Length 1053;
Best Local Similarity 70.9%; Pred. No. 7.5e-13;
Matches 166; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 286 GCCTTGAGCGGGAGCTGGAGCGGCTCGGGCCGAGCTGCGGGAGGAGCGGCAAGGCCAT 345
DB 187 GCCTTGCGCGGAGGTGGGGCGCTGCAGCCGAGCTGCGGCTGAGCGGCGGCGG 246
QY 346 GACCAGATGTCTCGGGCTCCACCATCAGCGGCTCGTGTGGAAGGAGGAGGAGAGAG 405
DB 247 GAGCGCAGGGTGCAGCTTCGCCGAGGAGCGCGGCTGTGGTGGAGGAGGAGAGAG 306
QY 406 GTGATTAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTATACGAGGAACCA 465
```

Db	307	GTGATCGAGTACGAGACGAGCTGCAGCTGAGCTGAGTGTGTACCGGCCAACCGAG	366
QY	466	CGCCTCGGAGAAGGCCCTTCGACGAGCTGGCAGCTGGGGACACGCGCGGGGAGCCCC	519
Db	367	CACCTGGAGCGCAAGCTCGCGGAGCGCGGGCGCAGGGGTGCAACGACGCCCC	420
RESULT 7			
LOCUS	CG385903	918 bp mRNA linear	EST 12-MAR-2001
DEFINITION	602454451F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4582669 5',		
ACCESSION	CG385903	mRNA sequence.	
VERSION	CG385903.1	GI:13279349	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://imgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-k@mail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: NIH Intramural Sequencing Center		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLCMI305 row: f column: 14		
	High quality sequence stop: 718.		
FEATURES	Location/Qualifiers		
source	1..918		
	/organism="Homo sapiens"		
	/db xref="taxon:9606"		
	/clone="IMAGE:4582669"		
	/clone_lib="NIH_MGC_15"		
	/tissue_type="adenocarcinoma cell line"		
	/lab host="DH10B (phage-resistant)"		
	/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:		
	ECORI; cDNA made by oligo-dT priming. Directionally		
	cloned into EcoRI/XhoI sites using the following 5'		
	adaptor: GGACGAG(G). Size-selected >500bp for average		
	insert size 1.8kb. Library constructed by Ling Hong in		
	the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies)"		
BASE COUNT	177 a 267 c 323 g 151 t		
ORIGIN			
Query Match	19.0%;	Score 121.2;	DB 12; Length 918;
Best Local Similarity	59.5%;	Pred. No. 3.8e-12;	
Matches	226;	Conservative 0;	Mismatches 148; Indels 6; Gaps 1;
QY	250	GACATGGGGCGGCCACCTTCCCGAGGAGCTGCCTCCTTCGACGGGAGCTGGAGCGG	309
Db	113	GCCAAAGTCAGCGGGCAGCAGCGCGGGTGTGGGGCCAGCTTGGGGGCCAGGTGGAGCGA	172
QY	310	CTSCGGGCCAGCTCGGGAGGAGCGGCAAGGCCATGACCATGTCTCTCGGGCTCCAG	369
Db	173	TTCCGGGTGAGCTGCAGCGGGAGCGCGCGGGGTGAGGAGCAGCGGGACAGCTTTGAG	232
QY	370	CATGAGCGGCTCGTGTGGAGGAGGAGGAGGAAAGGTGATTTCAGTACCAAGAAACAGCTG	429
Db	233	GGGAGCGGCTGCCTGGCAGGAGGAGGAGGAGGAGGAGCTGATCGCTACCAAGAGCAGCTG	292
QY	430	CAGCAGAGCTTACGTGGCCATGTATACCAAGCGAAACAGCGCTGGAGAGGCGCTTCAGCAG	489
Db	293	CAGCAACATACATCCAGATGTATCCGCGCGCAACCGGAGCTTACAGAGGAGCTGCAGCAG	352


```

High quality sequence stop: 589.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4130469"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab host="DHI0B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 199 a 247 c 311 g 144 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 901;
Best Local Similarity 100.0%; Pred. No. 6.2e-12;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 579
DB 3 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 62
QY 580 TGAGGGCTGCTGGGAGGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 639
DB 63 TGAGGGCTGCTGGGAGGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 122

RESULT 11
BG035858 926 bp mRNA linear EST 24-JAN-2001
LOCUS 602325955F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4414260 5',
DEFINITION mRNA sequence.
ACCESSION BG035858
VERSION BG035858.1 GI:12430413
KEYWORDS EST...
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 926)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM0141 row: m column: 13
High quality sequence stop: 735.
Location/Qualifiers
1..926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4414260"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 199 a 247 c 311 g 144 t
ORIGIN

High quality sequence stop: 589.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4130469"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab host="DHI0B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 199 a 247 c 311 g 144 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 901;
Best Local Similarity 100.0%; Pred. No. 6.2e-12;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 579
DB 3 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 62
QY 580 TGAGGGCTGCTGGGAGGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 639
DB 63 TGAGGGCTGCTGGGAGGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 122

RESULT 12
BG170653 957 bp mRNA linear EST 06-FEB-2001
LOCUS 602325232F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426924 5',
DEFINITION mRNA sequence.
ACCESSION BG170653
VERSION BG170653.1 GI:12677356
KEYWORDS EST...
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 957)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10174 row: m column: 05
High quality sequence stop: 681.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4426924"
/clone_lib="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 194 a 276 c 326 g 161 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 926;
Best Local Similarity 61.7%; Pred. No. 6.2e-12;
Matches 211; Conservative 0; Mismatches 125; Indels 6; Gaps 1;

QY 288 CCTCAGCGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGAGGCGCATGA 347
DB 5 CTTGCGGCGCCAGCTGGAGCGATTGCGGGTGAGCTGCAGCGGAGCGCGCGGGTGA 64
QY 348 CCAGATGTCCTCGGGCTTCAGCATGAGCGGCTGCTGTGGAAGGAGGAGGAGAGGT 407
DB 65 GGAGCAGCGGACAGCTTTGAGGGGGAGCGGCTGCGCTGGCAGGAGGAGGAGGAGGT 124
QY 408 GATTGAGTACAGAAACAGCTGCAGCAGAGTACGTGGCCATGTACCAAGCGAACCAGCG 467
DB 125 GATCCGCTACCAAGAGAGCTGCAGCACTATCAGATGTACCGCGCAACCGCA 184
QY 468 CCTGGAAGGCGCTGCAGCAGCTGGCACGTGGGGACAGCGCGGGGAGCGCTTGGAGGT 527
DB 185 GCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTGGAGCGCGGAGCTCGCTGACCT 244
QY 528 TGACCTGG-----AAGGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTG 581
DB 245 GGGCTGCGCGGAGCAGGCGCCCTGCTGCTGCTGGAGGAGATCACTGCTACTGAGATCTA 304
QY 582 AGGGGCTGCTGGGAGGCGAGTCTGGGGACCTGGCACTGGG 623
DB 305 GGGCGCTCAGCAACACGAGCTCTGTAGGAGCTCTGCCAGAGGG 346

FEATURES
source
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4426924"
/clone_lib="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 194 a 276 c 326 g 161 t
ORIGIN

```


ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 957;
 Best Local Similarity 61.7%; Pred. No. 6.2e-12;
 Matches 211; Conservative 0; Mismatches 125; Indels 6; Gaps 1;
 QY 288 CTTGAGCGGAGCTGGAGCGGCTCGCGGCGAGCTGCGGAGGAGCGGCAAGGCCATGA 347
 Db 8 CTTGCGGCGCCAGGTGGAGCGATTGCGGGTGGAGCTGACGGGAGCGCGCGGGGTGA 67
 QY 348 CCAGATGCTCTGGGCTTCCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAAGGT 407
 Db 68 GGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGCGCTGGCAGGAGGAGGAGGAGGT 127
 QY 408 GATTGAGTACAGAAACAGCTGCAGCAGAGCTAGCTGCGCCATGTACAGCGGAGACCGG 467
 Db 128 GATCGCTTACAGAAAGCAGCTGCAGCAACATACATCCAGATGTATCCGCGCAACCGGCA 187
 QY 468 CTTGAGAAAGGCGCTGCAGCAGCTGGCAGCTGGGAGACAGCGCGGGGAGCGCTTGGAGGT 527
 Db 188 GCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTGGAGCGCGGAGCTGCTGACCT 247
 QY 528 TGACCTGG-----AAGGGGTGACATCCCTACGAGACATCATAGCCATGAGATCTG 581
 Db 248 GGGCTGCGCGAGCAGGCGCCCTGCTGCTGCTGAGGAGATCACTGCTACTGAGATCTA 307
 QY 582 AGGGCTCGCTGGGAGCGAGCTCTGGGAGCCTGCGACTGGG 623
 Db 308 GGGCGCTCAGCAACAGCTCTGTAGGAGCTCTGCCAGAGGG 349

RESULT 13

LOCUS BG750395 923 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602709229F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4845799 5',
 mRNA sequence.

ACCESSION

VERSION BG750395.1 GI:14061048

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 923)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

plate: LCM1685 row: j column: 08

High quality sequence stop: 895.

Location/Qualifiers

1..923

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4845799"

/clone_lib="NIH_MGC_43"

/tissue_type="normal pigmented retinal epithelium"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally-

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library. |"

BASE COUNT 180 a 246 c 368 g 129 t

ORIGIN

Query Match 18.6%; Score 118.6; DB 12; Length 923;
 Best Local Similarity 52.8%; Pred. No. 1.1e-11;
 Matches 307; Conservative 0; Mismatches 239; Indels 35; Gaps 1;

QY 1 GAGTGAAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGCACACGCGG 60
 Db 269 GAGCTGGTGCAGAAAGGCGCAGCTGGTGGCTCTGCGGGTGGCGCTGCGGAGGCGCGT 328
 QY 61 GGCAAGCTGGAGGCGCTGGAGCTCAGGACCCAGACACCTGGAGGCGCCCTCGCACCAAG 120
 Db 329 GCTACGCTGCGGTCTAGTAGGCGCGCTGCGCGGGTCTACAGGAGCGCCCGGAGCTCGG 388
 QY 121 GGCCTGAGAGCTGGAGGCTGTGTAGAATGAGCTGCAGCGCAAGAAAGAACAGCGGAGCTG 180
 Db 389 GAGCTGAGAGCTGGAAAGCTGTTCACAGAGCTGCAGCGCACCCGCCAGGAAGCTGAGCAG 448
 QY 181 CTCGGGAGAGAGTGAACCTGCTGAGCAGAGAGCTGCAGGAGCTGCGGGCCCA----- 233
 Db 449 CTGCGGAGAGAAAGCTGGGCACTTGGATGCTGAGCGCGGAGCTCCGGGAGCCCTCTGTGC 508
 QY 234 -----GGCGGCGCTGGCGCGACATGGGCGCGCCCA 265
 Db 509 CACCTGCCACCGTTGACCCATTCTCTGCGCAGAGAGTGATGAGGCCAAAGTCAGCGGG 568
 QY 266 CTTTCCCGAGGACGTCCTGCGCTGCAGCGGGAGCTGGAGCGGCTGCGGCGCGAGCTGC 325
 Db 569 CAGCAGCGCGGCTTGGGGGCGAGCTTGGCGGCGCCAGCTGGAGCGATTTGGCGGTGGAGCTGC 628
 QY 326 GGGAGGAGCGCAAGGCCATGACCAGATGCTCTCGGGCTTCAGCATGAGCGGCTCGTGT 385
 Db 629 AGCGGAGCGCGCGCGGGTGAGCAGCAGCGCGGACAGCTTTGAGGGGAGCGGCTGGGCT 688
 QY 386 GGAAGGAGGAGAGGAGAGGTTGATTACGTACCAAGAACAGCTGCAGCAGAGCTACGTGG 445
 Db 689 GGCAGGAGAGAGAGAGAGGTTGATTACGTACCAAGAACAGCTGCAGCAGACACTACATCC 748
 QY 446 CCATGTACCGAGAACCAACAGCGCTTGGAGAGGCGCTTCAGCAGAGCTGCACAGCTGGGGACA 505
 Db 749 AGATGTACCGCGCAACCGGCGCAGCTAGAGCAGGAGCTGAGCAGCTCAGCTGAGGAGCTTG 808
 QY 506 GCGCGGGGAGCGCTTGGAGGTTGACCTGGAGGGGCTGAC 546
 Db 809 GAGGCGCGGAGCTCGTTGACCTGGGCTTGGCGCGAGCAGGC 849

RESULT 14

BM193264/c

LOCUS

DEFINITION

TCBAP1E6182 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP6182, mRNA
 sequence.

ACCESSION

BM193264

VERSION

BM193264.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 469)

REFERENCE

AUTHORS

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunarathne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished (2001)

CONTACT: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

FEATURES

source

FEATURES
source
Email: clones@txccc.org
Seq primer: M13 primer:
Location/Qualifiers
1. 469
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clones="TCBAP182"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/note="vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCAGCGCGCAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGGCTCGGATCCGCGCGCCGAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
a 175 c 132 g 99 c

BASE COUNT
ORIGIN

Query Match 18.5%; Score 118.4; DB 13; Length 469;
Best Local Similarity 59.4%; Pred. No. 1.1e-11;
Matches 222; Conservative 0; Mismatches 146; Indels 6; Gaps 1;
Qy 256 GGGCGGCCCTTCCCGGAGAGCTCCCTCCCTGCGAGCGGAGCTGGAGCGGCTGCGG 315
Db 455 GTGACGCGGCGAGCAGCGCGGGTTGGGGCGAGCTTGGCGGCCAGGTGGAGCGATTGCGG 396
Qy 316 GCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACAGATGCTCTCGGGCTTCCAGCATGAG 375
Db 395 GTGAGCTGACGGAAGCGCGGGTGGAGAGCGCGGAGCAGCTTTGAGGGGAG 336
Qy 376 CGGCTCGTGTGGAGGAGGAGAGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAG 435
Db 335 CGGCTGGCTTGGAGGAGAGAGAGAGGTGATCCGCTACAGAGCAGCTGCAGCAG 276
Qy 436 AGCTACGTGGCCATGTACACGGGAAACAGCGCTGGAGAGGCGCTTCGACAGCTGGCA 495
Db 275 AACTACATCCAGATGTACCGCGCAACCGGCGAGCTAGAGCAGGAGCTGCAGCAGCTCAGC 216
Qy 496 CGTGGGACAGCGCGGGAGCCCTTGGAGCTTGACCTGG-----AAGGGCTGACATC 549
Db 215 CTGAGCTGGAGGCGCGGAGCTCGCTGACCTGGGCTGGCGAGCGGCCCTTGATC 156
Qy 550 CCCTACGAGGACATCATGACCTGAGATCTGAGGGGCTGCTGGGAGGCGAGTCTGGG 609
Db 155 TGCCTGGAGGAGATCACTGCTACTGAGATCTAAGCCCTCAGCAACAGCTCTGTAGGA 96
Qy 610 GACTTGGCACTGGG 623
Db 95 GCTCTGCCAGAGGG 82

RESULT 15
BQ694854
LOCUS
DEFINITION
1000997 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
450 bp mRNA linear EST 15-JUL-2002
CDNA 5', mRNA sequence.
BQ694854
ACCESSION
BQ694854.1 GI:21820170
KEYWORDS
EST.

SOURCE
ORGANISM

human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 450)

AUTHORS

Yang, R.-Z., Shuldiner, A. and Gong, D.-W.

TITLE

EST analysis of human adipose gene expression

JOURNAL

Unpublished (2002)

COMMENT

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, HH497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGGAAGCGCGCATTTGTGTGGT

BACKWARD: AATACGACTCATATAGGCGGATTCG

Seq primer: GTTGTACCCGGGAATTC

Location/Qualifiers

1. 450

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

/sex="Male and Female"

/tissue_type="Adipose"

/note="Vector: lambdaTriplex"

BASE COUNT 86 a 131 c 172 g 61 t

ORIGIN

Query Match 18.5%; Score 118.2; DB 14; Length 450;
Best Local Similarity 61.2%; Pred. No. 1.1e-11;
Matches 211; Conservative 0; Mismatches 128; Indels 6; Gaps 1;
Qy 250 GACATGGGCGGCCACCTTCCCGAGAGCTCCCTCCCTGCGAGCGGAGCTGGAGCGG 309
Db 91 GCCAAAGTGCAGCGGCGAGCGCGGGTTGGGGCGAGCTTGGCGGCCAGGTGGAGCGA 150
Qy 310 CTGCGGCGCCAGCTGCGGGAGGAGCGGCAAGGCCATGACAGATGTCTCTCGGGCTTCCAG 369
Db 151 TTGCGGTGGAGCTGCAGCGGAGCGCGGGTGGAGGAGCAGCGGAGCAGCTTTGAG 210
Qy 370 CATGAGCGGCTCGTGTGGAGGAGGAGAGAGGTGATTTCAGTACCAGAAACAGCTG 429
Db 211 GGGAGCGGCTGGCTTGGCAGGAGAGAGAGGAGGAGGTGATCCGCTACCAAGCAGCTG 270
Qy 430 CAGCAGAGCTACGTGGCCATGTACAGCGGAGCAGCGCTGGAGAGGCCCTGCAGCAG 489
Db 271 CAGCACAACTACATCCAGATGTACCGGCGCAACCGGAGCTAGAGCAGGAGCTGCAGCAG 330
Qy 490 CTGGCAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGG-----AAGGGGCT 543
Db 331 CTGAGCTGGAGCTGGAGGCGCGGAGCTCGTGACTGGGCTTGGCGGAGCGGCCCC 390
Qy 544 GACATCCCTTACGAGGACATCATGCCACTGAGATCTGAGGGGCT 588
Db 391 TGCATCTGCTGGAGGAGATCACTGCTACTGAGATCTAGGGCCCT 435

Search completed: June 15, 2003, 03:41:57
Job time : 894.672 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 4660.9 Seconds
(without alignments)
11183.052 Million cell updates/sec

Title: US-09-513-888c-3
Perfect score: 1791
Sequence: 1 atgggcagcgtcagtagcct.....tcatagcactgagatctga 1791

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.in.*

18: em.mu.*

19: em.om.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

28: em.vi.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1791	100.0	5492	9	AF123659 Homo sapi
2	1643	91.7	1722	9	AF123658 Homo sapi
3	1583	88.4	1692	9	AF123657 Homo sapi
4	1427	79.7	1614	9	AF123655 Homo sapi
5	1372.4	76.6	1515	9	AF123656 Homo sapi
6	808.8	45.2	173264	2	AC025853 Homo sapi
7	807.2	45.1	9108	9	AF123653 Homo sapi
8	592.6	33.1	227884	2	AC114995 Mus muscu
9	592.6	33.1	263546	2	AC099416 Mus muscu
10	582.6	31.4	633	9	AF123654 Homo sapi
11	475.4	26.5	191210	2	AC108987 Rattus no
12	314	17.5	1935	9	AY029201 Homo sapi
13	314	17.5	2275	9	HSM805394
14	314	17.5	2766	9	BC006212 Homo sapi
15	314	17.5	5257	9	AB011124 Homo sapi
16	314	17.5	5733	9	AB058716 Homo sapi
17	310.8	17.4	2099	9	AB046013 Macaca fa
18	284.8	15.9	2767	10	BC014695 Mus muscu
19	278	15.5	1612	9	BC005855 Homo sapi
20	274	15.3	191210	2	AC108987 Rattus no
21	265.8	14.8	2416	10	RNO378801 Rattus no
22	218.4	12.2	85530	2	AC125907 Rattus no
23	197	11.0	181086	9	AL133215 Human DNA
24	189	10.6	155662	2	AC112388 Rattus no
25	186.4	10.4	272545	2	AC090533 Mus muscu
26	184.6	10.3	180418	2	AC106674 Rattus no
27	184.6	10.3	195673	2	AL807807 Mus muscu
28	181	10.1	168210	2	AC116700 Mus muscu
29	180.8	10.1	2308	9	AK097997 Homo sapi
30	179.4	10.0	82400	10	AC090495 Genomic s
31	168.8	9.4	103729	2	AC096620 3
32	168.8	9.4	270700	2	AC105485 Rattus no
33	168.6	9.4	167636	2	AC107097 Rattus no
34	167.6	9.4	207683	2	AC098712 Mus muscu
35	167.2	9.3	125856	9	HS1187M17 Human DNA
36	166	9.3	231	9	AF123652 Homo sapi
37	165.4	9.2	195690	10	AC098726 Mus muscu
38	159.8	8.9	177035	2	AC099443 Rattus no
39	157.4	8.8	141640	2	AC016204 Homo sapi
40	157.4	8.8	180553	2	AC011767 Homo sapi
41	156.8	8.8	198228	2	AC102097 Mus muscu
42	156	8.7	227588	2	AC094571 Rattus no
43	155.6	8.7	204259	2	AC110817 Mus muscu
44	152.6	8.5	184865	2	AL807824 Mus muscu
45	152.2	8.5	2508	9	M25312 Orang-utan

ALIGNMENTS

RESULT 1
AF123659
LOCUS AF123659 Homo sapiens FEZ1 (FEZ1) mRNA, complete cds. 5492 bp linear PRI 07-APR-1999
DEFINITION Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.
ACCESSION AF123659
VERSION AF123659.1 GI:4572475
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5492)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,


```
Db 1021 GAGAGCGGCGAGCTCCGACAGAGCTCGAGAGCTCATGAAGAGCAGGACCTGCTGGAG 1080
Qy 1081 ACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTTCGGCCCGCGCTGGAGAGACC 1140
Db 1081 ACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTTCGGCCCGCGCTGGAGAGACC 1140
Qy 1141 CAGTGGAGGTGTGCGACAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1200
Db 1141 CAGTGGAGGTGTGCGACAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1200
Qy 1201 TCCACAGCGAGGTGAAGCCAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1260
Db 1201 TCCACAGCGAGGTGAAGCCAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1260
Qy 1261 GACACGCGGGCAAGCTGAGAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1320
Db 1261 GACACGCGGGCAAGCTGAGAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1320
Qy 1321 CGACACAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1380
Db 1321 CGACACAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1380
Qy 1381 CGGAGAGTCTGCGGAGAGTGAAGCTGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1440
Db 1381 CGGAGAGTCTGCGGAGAGTGAAGCTGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1440
Qy 1441 CAGGCGGCGCTGCGGCGGAGTGAAGCTGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1500
Db 1441 CAGGCGGCGCTGCGGCGGAGTGAAGCTGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1500
Qy 1501 CAGGCGGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1560
Db 1501 CAGGCGGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1560
Qy 1561 ATGCTCGGCTTCCAGCATGAGCGCTCGTGTGAAGAGGAGAGAGGAGAGAGTGAAT 1620
Db 1513 -----GAGCGGCTGTGTGAAGAGGAGAGAGGAGAGTGAAT 1551
Qy 1621 CAGTACAGAGAACAGCTCAGAGAGTCAAGTGTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1680
Db 1552 CAGTACAGAGAACAGCTCAGAGAGTCAAGTGTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1611
Qy 1681 GAGAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1740
Db 1612 GAGAGGCGCTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1671
Qy 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGGACCTGAGATCTGA 1791
Db 1672 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGGACCTGAGATCTGA 1722
```

```
RESULT 3
AF123657
LOCUS AF123657
DEFINITION Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
complete cds.
ACCESSION AF123657
VERSION AF123657.1 GI:4572471
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1692)
AUTHORS Ishii.H., Baffa.R., Numata.S.I., Murakumo.Y., Rattan.S., Inoue.H.,
Mori.M., Fidanza.V., Alder.H. and Croce.C.M.
The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 1692)
```

AUTHORS
TITLE
JOURNAL
FEATURES
source

Ishii.H., Baffa.R., Numata.S.I., Murakumo.Y., Rattan.S., Inoue.H.,
Mori.M., Fidanza.V., Alder.H. and Croce.C.M.
Direct Submission
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA
Location/Qualifiers
1..1692

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="G3611"
1..1692
/gene="FEZ1"
1..1692

gene
CDS

/gene="FEZ1"
/note="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23838.1"
/db_xref="GI:4572472"
1..1692

/translation="MGSVSSLSISGHSFKHCRASQYKLRKSHLKLNRYSGLLRF
GFSQDSGHGKSSKMGKSEDFYIKVSKARGSHHPDYTALSSDGLGQAGVDFDPST
PKMLPSPNQLMEGSKGAVPTAFKPSVSSVOLDPLVTPVPTSRFGSAHNIQGV
ELKPGLCGALSDSGRNMSLLPHTSPVSSVOLDPLVTPVPTSRFGSAHNIQGV
LQDSNMSLXALSFSDGSKLHSHKADKGPSVRSPISTDECSIOLEOKLREGA
LQKQSFESKELASLAYERPRRCDELEGPEPKGNLKNQSKQSQRAQVHLIQ
VLQDQEKRLQFELSMEKQDLLETKLSYREKTSFPALESTQWECQKSGEIS
LLKQKESQTEVNAKASEILGKAQKDRKGLGLELRTQDLGLARTKGLLEVC
ENELQKKNABEILREKVNLLERAEIREROCHDOMSSGFQHERLVWKEKEKVIQ
YQKQLOSVYVYQYRNORLEKALQQLARGDSAGEPLEVDLEGADIPYEDIATEI"
BASE COUNT 380 a 522 c 554 g 236 t

Query Match 88.4%; Score 1583; DB 9; Length 1692;
Best Local Similarity 94.5%; Pred. No. 3.6e-214;
Matches 1692; Conservative 0; Mismatches 0; Indels 99; Gaps 1;

```
Qy 1 ATGGGCGAGCTCAGTACGCTCATCTCGGCGCACAGCTTCCACAGCAAGCACTGCGGGCT 60
Db 1 ATGGGCGAGCTCAGTACGCTCATCTCGGCGCACAGCTTCCACAGCAAGCACTGCGGGCT 60
Qy 61 TCGCAGTACAGCTGCGCAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
Db 61 TCGCAGTACAGCTGCGCAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
Qy 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAACGGCAAGTCCAGCTCAAAATGGGC 180
Db 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAACGGCAAGTCCAGCTCAAAATGGGC 180
Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCA 240
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCA 240
Qy 241 GATTACAGGCACTGTCCAGCGGGGATTAGGGGGGCGAGGCTGGGGTGGACTTTGACCCG 300
Db 241 GATTACAGGCACTGTCCAGCGGGGATTAGGGGGGCGAGGCTGGGGTGGACTTTGACCCG 300
Qy 301 TCCACACCCCAAGCTCATGCCCCCTTCTCCAAATCAGCTAGAAAATGGGTCGAGAAAGGT 360
Db 301 TCCACACCCCAAGCTCATGCCCCCTTCTCCAAATCAGCTAGAAAATGGGTCGAGAAAGGT 360
Qy 361 GCAGTAGGGCCACAGCTTCAAGCTGTGCTGCACGGTCAAGGAGCCATCTGCACCTCC 420
Db 361 GCAGTAGGGCCACAGCTTCAAGCTGTGCTGCACGGTCAAGGAGCCATCTGCACCTCC 420
Qy 421 TCCCGGAGAGTGCAGCCACAGCTGCACCCGCCCTCTCCAGACAGCCCAAGGAGCAG 480
Db 421 TCCCGGAGAGTGCAGCCACAGCTGCACCCGCCCTCTCCAGACAGCCCAAGGAGCAG 480
Qy 481 GAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGACACTCCGGCCCGGAACTCCATGTCC 540
Db 481 GAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGACACTCCGGCCCGGAACTCCATGTCC 540
```

```
QY 541 AGCTGCCCCACACAGCAGCAGCAGCAGCTACAGCTGGACCCCGCTGGTCAACCCGCTG 600
Db 541 AGCTGCCCCACACAGCAGCAGCAGCAGCTACAGCTGGACCCCGCTGGTCAACCCGCTG 600
QY 601 GGACCCCAAGCGCTTTTGGGGCTTCGCCACACATCACCAGGCGATCGTCTCCAG 660
Db 601 GGACCCCAAGCGCTTTTGGGGCTTCGCCACACATCACCAGGCGATCGTCTCCAG 660
QY 661 GACAGCAACATGAGCTGAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGC 720
Db 661 GACAGCAACATGAGCTGAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGC 720
QY 721 CACTCGAACAAGCGACAGAGGCGCTCGTGTGTCTCCCTCCCCCATCTCCACGGACAG 780
Db 721 CACTCGAACAAGCGACAGAGGCGCTCGTGTGTCTCCCTCCCCCATCTCCACGGACAG 780
QY 781 TGCAGCATCCAGGAGCTGAGCAGAGCTGTGGAGAGGAGGCGCCCTCCAGAGCTG 840
Db 781 TGCAGCATCCAGGAGCTGAGCAGAGCTGTGGAGAGGAGGCGCCCTCCAGAGCTG 840
QY 841 CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCTACGAGGAGCGCGCGG 900
Db 841 CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCTACGAGGAGCGCGCGG 900
QY 901 CGCTGACGGACGAGCTGAGGCGCGGAGCCCAAGGCGGCAACAGCTCAAGCAGGCC 960
Db 901 CGCTGACGGACGAGCTGAGGCGCGGAGCCCAAGGCGGCAACAGCTCAAGCAGGCC 960
QY 961 TCGCAGAGAGCAGCGCGCAGCAGCTCTGCACCTGAGGAGTCTGAGCTTCAAGCAG 1020
Db 961 TCGCAGAGAGCAGCGCGCAGCAGCTCTGCACCTGAGGAGTCTGAGCTTCAAGCAG 1020
QY 1021 GAGAAGCGGAGCTCCGCGAGAGCTCGAGAGCTCATGAAGGAGCAGGACTGCTGAG 1080
Db 1021 GAGAAGCGGAGCTCCGCGAGAGCTCGAGAGCTCATGAAGGAGCAGGACTGCTGAG 1080
QY 1081 ACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCCCGCTGGAGAGACC 1140
Db 1081 ACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCCCGCTGGAGAGACC 1140
QY 1141 CAGTGGGAGGTGTCAGAGCTCAGGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200
Db 1141 CAGTGGGAGGTGTCAGAGCTCAGGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200
QY 1201 TCCAGACGAGGTGAACCAAGCTAGCAGATCTCGGTCTCAAGGCGACAGCTGAAG 1260
Db 1201 TCCAGACGAGGTGAACCAAGCTAGCAGATCTCGGTCTCAAGGCGACAGCTGAAG 1260
QY 1261 GACACGCGGGCAAGCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1320
Db 1261 GACACGCGGGCAAGCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1320
QY 1321 CGCACCAAGGCGCTGAGCTGAGGTCTGTGAGATGAGCTGAGCGGCAAGAGAACGAG 1380
Db 1321 CGCACCAAGGCGCTGAGCTGAGGTCTGTGAGATGAGCTGAGCGGCAAGAGAACGAG 1380
QY 1381 CGCGAGCTGCTGCGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGCC 1440
Db 1381 CGCGAGCTGCTGCGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGCC 1440
QY 1441 CAGGCGCGCTTGGCGCGGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1500
Db 1441 CAGGCGCGCTTGGCGCGGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1500
QY 1501 CAGCGGAGCTGAGAGCGCTCGGCGCGAGCTGCGGAGGAGCGGAGCCATGACCGAG 1560
Db 1501 CAGCGGAGCTGAGAGCGCTCGGCGCGAGCTGCGGAGGAGCGGAGCCATGACCGAG 1560
QY 1561 ATGCTCTGGGCTTCCAGCATGAGCGCTCTGTGGAGGAGGAGGAGGAGGAGGAGTATT 1620
Db 1561 ATGCTCTGGGCTTCCAGCATGAGCGCTCTGTGGAGGAGGAGGAGGAGGAGGAGTATT 1620
QY 1621 ATGCTCTGGGCTTCCAGCATGAGCGCTCTGTGGAGGAGGAGGAGGAGGAGGAGTATT 1680
Db 1621 ATGCTCTGGGCTTCCAGCATGAGCGCTCTGTGGAGGAGGAGGAGGAGGAGGAGTATT 1680
```

```
QY 1621 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGCCCATGTACCGAGGAAACAGCGCTG 1680
Db 1522 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGCCCATGTACCGAGGAAACAGCGCTG 1581
QY 1681 GAGAAGGCCCTTCAGCAGCTGCGCAGCTGGGACAGCGCGGGAGCCCTTGGAGTTGAC 1740
Db 1582 GAGAAGGCCCTTCAGCAGCTGCGCAGCTGGGACAGCGCGGGAGCCCTTGGAGTTGAC 1641
QY 1741 CTGGAAGGGGCTGACATCCCTACGAGGACATCATAGCCTGAGATCTGA 1791
Db 1642 CTGGAAGGGGCTGACATCCCTACGAGGACATCATAGCCTGAGATCTGA 1692

RESULT 4
AF123655
LOCUS Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, alternatively
DEFINITION spliced, complete cds.
ACCESSION AF123655
VERSION AF123655.1 GI:4572467
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1614)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 1614)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES
Location/Qualifiers
1..1614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="T8D145M4"
1..1614
/gene="FEZ1"
1..1614
/gene="FEZ1"
/notes="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23836.1"
/db_xref="GI:4572468"
/translation="MGSVSSLISGSHPSKHCASOYKLRKSHLKLNRYSDDLRF
GFSDSGHGKSSKMGSEDFYIKVSKARGSHHPDYTALSSDLGSGOAGVDFDPS
PKLMPFNSOLEMGSEKAVRPTAFKVLPRSGAILHSSPESASHQHPAPDPKPEQ
ELKPLGCSALSDSGRNSSLPHTSSSTQLDPLVTPVGPTRFSGSAHNTTQGI
LQDSNMMSLKALSFSDGSKLGHNSKADGKPCVRSPISTDECSIQLEQKLLREGA
LQKRSFBEKELASLAYERPRRCDELEGPEKPGKNLQKASQKSAQQVHLQ
VLQOQKRLROBLESLMKEDLLETKLARSYREKTSFGPALSETOWECQKSGEIS
LLKQOLKESQTEVNAKASEILGLKAOLKDTGKLEGLRTQDLEGALRTKGLELVC
ENELORKNAEALLREKHLRWKEKEKVIQYQKLOQSQSYVAMYQRNLEKALQQL
ARGDSAGEPLEVDLEGADIPYDIIATEI"
BASE COUNT 367 a 500 c 522 g 225 t
ORIGIN

Query Match 79.7%; Score 1427; DB 9; Length 1614;
Best Local Similarity 90.1%; Pred. No. 3.7e-192;
Matches 1614; Conservative 0; Mismatches 0; Indels 177; Gaps 1;
QY 1 ATGGGCAGCGCTCAGTGCCTCATCTCCGGCCACACAGCTTCCACAGCAAGCAAGCACTGCCGGGCT 60
```

Db	1	ATGGCGAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTCCCGGGCT	60
Qy	61	TCGCAGTACAAAGCTGCCAAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG	120
Db	61	TCGCAGTACAAAGCTGCCAAAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG	120
Qy	121	CTGCTGAGGTTTGCTTCTCCAGGACTCCGGTCACGGCAAGTCCAGTCCAAAATGGGC	180
Db	121	CTGCTGAGGTTTGCTTCTCCAGGACTCCGGTCACGGCAAGTCCAGTCCAAAATGGGC	180
Qy	181	AAGAGCGAAGACTTCTTCTATCAAGGTGAGCCAGAAAGCCCGGGGCTCCCATCAACCA	240
Db	181	AAGAGCGAAGACTTCTTCTATCAAGGTGAGCCAGAAAGCCCGGGGCTCCCATCAACCA	240
Qy	241	GATTACACGGCACTGTCACGGGGGATTTAGGGGCCAGAGTGGGTGGAGCTTTGACCCG	300
Db	241	GATTACACGGCACTGTCCAGCGGGATTTTAGGGGCCAGAGTGGGTGGAGCTTTGACCCG	300
Qy	301	TCCACACCCCAAGCTCATGCCCTTCTTCAATCAGCTAGAAATGGGCTCCGAGAAGGT	360
Db	301	TCCACACCCCAAGCTCATGCCCTTCTTCAATCAGCTAGAAATGGGCTCCGAGAAGGT	360
Qy	361	GCAGTGAGGCCACAGCCTTCAAGCTGTCTGTCACGGTCAGGAGCATCTCTGCACTCC	420
Db	361	GCAGTGAGGCCACAGCCTTCAAGCTGTCTGTCACGGTCAGGAGCATCTCTGCACTCC	420
Qy	421	TCCCGGAGAGTGCCAGCACCAAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAG	480
Db	421	TCCCGGAGAGTGCCAGCACCAAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAG	480
Qy	481	GAGCTGAAGCCTGSCCTGTGCTCTGGGGCGCTGTGACACTCCGGCCGGAACCTCATGTCC	540
Db	481	GAGCTGAAGCCTGSCCTGTGCTCTGGGGCGCTGTGACACTCCGGCCGGAACCTCATGTCC	540
Qy	541	AGCCTGCCACACACAGCACACAGCAGACTACAGCTGGAACCCGTGGTTCACACCCGTG	600
Db	541	AGCCTGCCACACACAGCACACAGCAGACTACAGCTGGAACCCGTGGTTCACACCCGTG	600
Qy	601	GGACCCACAAGCCGTTTTGGGGGCTCCGCCACAACATACCCAGGAGCATGCTCTCCAG	660
Db	601	GGACCCACAAGCCGTTTTGGGGGCTCCGCCACAACATACCCAGGAGCATGCTCTCCAG	660
Qy	661	GACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGC	720
Db	661	GACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGC	720
Qy	721	CACCTCGAACAGGCGACAAAGGGCCCTCGTGTGTCCGCTCCCCCATCTCCACGGACGAG	780
Db	721	CACCTCGAACAGGCGACAAAGGGCCCTCGTGTGTCCGCTCCCCCATCTCCACGGACGAG	780
Qy	781	TGCAGCATCCAGGACTGGAGCAAGAGTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG	840
Db	781	TGCAGCATCCAGGACTGGAGCAAGAGTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG	840
Qy	841	CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTACGAGGAGCGCCGCGG	900
Db	841	CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTACGAGGAGCGCCGCGG	900
Qy	901	CGCTGCAGGGACGAGCTGGAGGGCCCGAGCCCAAGCGCGCAACAAAGCTCAAGCAGGCC	960
Db	901	CGCTGCAGGGACGAGCTGGAGGGCCCGAGCCCAAGCGCGCAACAAAGCTCAAGCAGGCC	960
Qy	961	TCGCAAGAGACGACGCGCGCCAGAGTCTCTGCACTGCAAGGTATCTGCAAGCTTACAGCAG	1020
Db	961	TCGCAAGAGACGACGCGCGCCAGAGTCTCTGCACTGCAAGGTATCTGCAAGCTTACAGCAG	1020
Qy	1021	GAGAAAGCGGACGCTCCGCGCAGAGCTCGAGAGCCTCATGAGGAGCAGGACCTGCTGGAG	1080
Db	1021	GAGAAAGCGGACGCTCCGCGCAGAGCTCGAGAGCCTCATGAGGAGCAGGACCTGCTGGAG	1080
Qy	1081	ACCAAGCTCAGGTCCTACGAGAGGGAGAGACCAAGCTTCGGGCCCGCGCTGGAGGAGACC	1140

Db	1081	ACCAAGCTCAGGTCCTTACGAGAGGGAGAA	GACACAGCTTCCGGCCCGCGCTGGAGGAGACC	1140
Qy	1141	CAGTGGGAGGTGTGCCAAGAAAGTCAGGCGAGATCTCCCTCTCTGAAGCAGCAGCTGAAGGAG	1200	
Db	1141	CAGTGGGAGGTGTGCCAAGAAAGTCAGGCGAGATCTCCCTCTCTGAAGCAGCAGCTGAAGGAG	1200	
Qy	1201	TCCCAGACGAGAGGTGAACGCCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACACGCTGAAG	1260	
Db	1201	TCCCAGACGAGAGGTGAACGCCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACACGCTGAAG	1260	
Qy	1261	GACACGCGGGCAAGCTGGAGGGGCTCGAGCTGAGGACCCAGGACCTTGGAGGGGCGCCCTG	1320	
Db	1261	GACACGCGGGCAAGCTGGAGGGGCTCGAGCTGAGGACCCAGGACCTTGGAGGGGCGCCCTG	1320	
Qy	1321	CGCACCAAGGGCTTGGAGCTGGAGGTCCTGTGAGAAATGAGCTGCAGCGCCAAAGAAACGAG	1380	
Db	1321	CGCACCAAGGGCTTGGAGCTGGAGGTCCTGTGAGAAATGAGCTGCAGCGCCAAAGAAACGAG	1380	
Qy	1381	GCAGAGCTGTCGGGAGAGAGGTGAACCTCTCTGGAGCAGGAGCTGCAGGAGCTCGCGGGCC	1440	
Db	1381	GCAGAGCTGTCGGGAGAGAGGTGAACCTCTCTGGAGCTGTGCGGGGAGA	1399	
Qy	1441	CAGGCCCGCTTGGCCCCGCGACATGGGGCCGCGCACCTTCCCGAGGACGTCCTCTGCCCTG	1500	
Db	1400	-----	-----	1399
Qy	1501	CAGCGGAGCTGGAGCGGCTTGGCGGCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACAG	1560	
Db	1400	-----	-----	1399
Qy	1561	ATGTCCTCGGCTTCCAGCATGAGCGCTCGTGTGGAAGAGGAGAAAGGAGAAGGTGATT	1620	
Db	1400	-----AGCATGAGCGCTCGTGTGGAAGAGGAGCAAGGAAGGTGATT	1443	
Qy	1621	CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAAGCGGAACACAGCGGCTG	1680	
Db	1444	CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAAGCGGAACACAGCGGCTG	1503	
Qy	1681	GAGAAGGCCCTTGCAGCAGCTGGCACGTGGGGACAGCCCGGGAGGCCCTTGGAGGTTGAC	1740	
Db	1504	GAGAAGGCCCTTGCAGCAGCTGGCACGTGGGGACAGCCCGGGAGGCCCTTGGAGGTTGAC	1563	
Qy	1741	CTGAAGGGGTGCATCCCCCTACGAGGACATCATAGCCACTGAGATCTGA	1791	
Db	1564	CTGAAGGGGTGCATCCCCCTACGAGGACATCATAGCCACTGAGATCTGA	1614	

RESULT 5	
AF123656	
LOCUS	linear mRNA PRI 07-APR-1999
DEFINITION	Homo sapiens clone D14 FEZ1 (FEZ1) mRNA, alternatively spliced,
ACCESSION	AF123656
VERSION	complete cds.
KEYWORDS	AF123656.1 GI:4572469
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1515)
AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M. The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999) 99199287
JOURNAL	
MEDLINE	10097140
PUBMED	2 (bases 1 to 1515)
REFERENCE	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M. Direct Submission
AUTHORS	
TITLE	Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107 USA
JOURNAL	

FEATURES	Location/Qualifiers	
	1..1515	
source	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="8"	
	/map="8p22"	
	/clone="D14"	
gene	1..1515	
	/gene="FEZ1"	
CDS	1..1515	
	/gene="FEZ1"	
	/note="alternatively spliced"	
	/codon_start=1	
	/product="FEZ1"	
	/protein_id="AAD23837.1"	
	/db_xref="GI:4572470"	
	/translation="MGSVSLISGHSFKHCRASOYKLRKSHLKLNRYSGLLRF	
	GFSODSGHKSSXMGKSEDFYIKVSKARGSHHPDYALSSDGLGGQAGVDFDPS	
	PPKLMPSNQJEMSEKCAVPTAFKVPVPRSGAILHSSPESASHQLHPAPDPKPEQ	
	ELKPLGSLGALSDGRMSMLPHTSTSSYOLDPLVTPVGFTRFGGSAHNTGGIV	
LDGNMMSLKALSDGSKLGNKADKPGSCVRSPISTDECSIQELQKLLREGA		
LQKQRFEEKELASSLAYERPRRCRDELEGPFGKNLQKQASQSORAQOVLHLQ		
VLQOQERKROLQBLESLMKPDLLLETKLRSYREKTSFGPALBETOWEVCOKSGEIS		
LLKQQLKESQTEVNAKASEIILGLKQQLKTRKLEGLFTQDLEGALRTKGLLEVC		
ENEIQQSVAMYQNRLEKALQQLARGDSAGEFLEVDLEGADIPYEDIATETI"		
BASE COUNT	336 a	483 c 482 g 214 t
ORIGIN		
	Query Match	76.6%; Score 1372.4; DB 9; Length 1515;
	Best Local Similarity	95.4%; Pred. No. 1.9e-184;
	Matches 1425; Conservative	0; Mismatches 66; Indels 3; Gaps 1;
QY	1	ATGGGACGGTCAGTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCCGGGCT 60
DB	1	ATGGGACGGTCAGTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCCGGGCT 60
QY	61	TCCGAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCGAGCGG 120
DB	61	TCCGAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCGAGCGG 120
QY	121	CTGCTGAGGTTGGCTTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGC 180
DB	121	CTGCTGAGGTTGGCTTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGC 180
QY	181	AAGACGGAAGACTTCTTCTACATCAAGGTACAGCTACGCAAGCCGGGCTCCCATCACCCA 240
DB	181	AAGACGGAAGACTTCTTCTACATCAAGGTACAGCTACGCAAGCCGGGCTCCCATCACCCA 240
QY	241	GATTACAGGCACTGTCAGCGGGGATTTAGGGGCCACAGCTGGGGTGGACTTTGACCCG 300
DB	241	GATTACAGGCACTGTCAGCGGGGATTTAGGGGCCACAGCTGGGGTGGACTTTGACCCG 300
QY	301	TCCACACCCCAAGCTATGCTCCCTTCTCCATCAGCTAGAAATGGGCTCCGAGAAGGT 360
DB	301	TCCACACCCCAAGCTATGCTCCCTTCTCCATCAGCTAGAAATGGGCTCCGAGAAGGT 360
QY	361	GCAGTGAGGCCACACAGCTTCAAGCTGTGTGTCACCGGTACAGAGCCATCTGCACTCC 420
DB	361	GCAGTGAGGCCACACAGCTTCAAGCTGTGTGTCACCGGTACAGAGCCATCTGCACTCC 420
QY	421	TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAG 480
DB	421	TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAG 480
QY	481	GAGCTGAAGCTGGCTGTGCTTGGGGGCTGTGCTGAGCTCCGGCCGGAATCCATGTCC 540
DB	481	GAGCTGAAGCTGGCTGTGCTTGGGGGCTGTGCTGAGCTCCGGCCGGAATCCATGTCC 540
QY	541	AGCTGCCACACACAGCAGCAGCTACAGCTGGAGCCCGCTGGTCCACACCCGTG 600
DB	541	AGCTGCCACACACAGCAGCAGCTACAGCTGGAGCCCGCTGGTCCACACCCGTG 600
QY	601	GGACCCACAAGCGGTTTTGGGGGCTCCGCCCAACAACATCACCCAGGGCATCGTCTCCAG 660

601	GGACCCACAAGCGGTTTTGGGGGCTCCGCCCAACAACATCACCCAGGGCATCGTCTCTCCAG	660
661	GACAGCAACATGATGAGCTTGAAGGCTCTGTCTTCTTCCGAGCGAGGTAGCAAGCTGGGC	720
661	GACAGCAACATGATGAGCTTGAAGGCTCTGTCTTCTTCCGAGCGAGGTAGCAAGCTGGGC	720
721	CATCTGAAACAAGGAGCAGCAAGGGCCCTCTGTGTGTCTCGCTCCCATCTCTCCACGAGCAG	780
721	CATCTGAAACAAGGAGCAGCAAGGGCCCTCTGTGTGTCTCGCTCCCATCTCTCCACGAGCAG	780
781	TGCAGCATCCAGGAGCTGGAGCACAAGCTGTGTGAGAGGAGGGCGCCCTCCAGAAAGCTG	840
781	TGCAGCATCCAGGAGCTGGAGCACAAGCTGTGTGAGAGGAGGGCGCCCTCCAGAAAGCTG	840
841	CAGCCACAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCTTGGCTTACAGGAGCGCCCGCGG	900
841	CAGCCACAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCTTGGCTTACAGGAGCGCCCGCGG	900
901	CGCTGCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC	960
901	CGCTGCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC	960
961	TCGAGAGAGCCAGCGCGGCGCAGCAGGTCTCTGACCTGCAGGTACTGCAGCTTCACGAG	1020
961	TCGAGAGAGCCAGCGCGGCGCAGCAGGTCTCTGACCTGCAGGTACTGCAGCTTCACGAG	1020
1021	GAGAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTGCTGCGAG	1080
1021	GAGAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTGCTGCGAG	1080
1081	ACCAAGCTCAGGTCCTACGAGAGGAGAAGACCAAGCTTCCGCCCGCGCTGGAGGAGACC	1140
1081	ACCAAGCTCAGGTCCTACGAGAGGAGAAGACCAAGCTTCCGCCCGCGCTGGAGGAGACC	1140
1141	CAGTGGAGGTGTCCAGAGTCAAGGCGAGATCTCCCTCTCTGAGAGCAGCAGCTGAAGAG	1200
1141	CAGTGGAGGTGTCCAGAGTCAAGGCGAGATCTCCCTCTCTGAGAGCAGCAGCTGAAGAG	1200
1201	TCCAGAGCGAGGTGAACGCCAAGCTAGCGAGATCTTGGGTCTCAAGGACACAGCTGAAG	1260
1201	TCCAGAGCGAGGTGAACGCCAAGCTAGCGAGATCTTGGGTCTCAAGGACACAGCTGAAG	1260
1261	GACACGCGGGCAAGCTGGAGGGCTCGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTG	1320
1261	GACACGCGGGCAAGCTGGAGGGCTCGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTG	1320
1321	CGACCAAGGGCTGGAGCTGGAGGTCTGTGAGATGAGTGCAGC---GCAAGAGAAGC	1377
1321	CGACCAAGGGCTGGAGCTGGAGGTCTGTGAGATGAGTGCAGC---GCAAGAGAAGC	1377
1378	GAGCGGAGGCTGTGCGGGAGAAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGG	1437
1381	ATGTACACAGCGAAGCAGCGCTGGAGAGCCCTGACAGAGCTGGACAGCTGGGAGCAGC	1440
1438	GCCACAGGCGCCCTGGCGGACATGGGGCGCCACCTTCCCGGAGGAGCTC	1491
1441	GCCGGGAGCCCTTGGAGGTTGACCTGGAAGGGGCTGCATCCCTACGAGGAC	1494

RESULT 6

AC025853

LOCUS

DEFINITION Homo sapiens chromosome 8 clone RP11-353K12 map 8, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.

ACCESSION AC025853

VERSION AC025853.13 GI:21431202

KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 173264)

AUTHORS JOURNAL REFERENCE AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-353K12
Unpublished
2 (bases 1 to 173264)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE JOURNAL REFERENCE AUTHORS

Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 173264)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL COMMENT

Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 17, 2002 this sequence version replaced gi:21321864.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7454
Center clone name: 353_K_12

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

FEATURES source

* as soon as it is available and the accession number will
* be preserved.
* 1 39461: contig of 39461 bp in length
* 39462 39561: gap of 100 bp
* 39562 108347: contig of 68786 bp in length
* 108348 108447: gap of 100 bp
* 108448 173264: contig of 64817 bp in length.

Location/Qualifiers

1. 173264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-353K12"
/clone_lib="RP11-11 Human Male BAC"
BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others

ORIGIN

Query Match 45.2%; Score 808.8; DB 2; Length 173264;
Best Local Similarity 99.8%; Pred. No. 2.7e-105;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 345 GGGCTCCGAGAGGGTGCAGTGGAGGCCACACAGCCTTCAAGCCTGTGTGTCACACGGTCCAGG 404
Db 23928 GGGCTCCGAGAGGGTGCAGTGGAGGCCACACAGCCTTCAAGCCTGTGTGTCACACGGTCCAGG 23987
Qy 405 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGA 464
Db 23988 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGA 24047
Qy 465 CAAGCCCAAGAGGAGGAGTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 524
Db 24048 CAAGCCCAAGAGGAGGAGTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 24107
Qy 525 CCGGAATCTCATGTTCAGCTGCGCCACACAGCAGCAGCAGCAGTACACAGCTGGACCC 584
Db 24108 CCGGAATCTCATGTTCAGCTGCGCCACACAGCAGCAGCAGTACACAGCTGGACCC 24167
Qy 585 GCTGGTCAACCCGCTGGAGCCCAAGCGCTTTTGGGGGCTCCGCCACACATCACCACA 644
Db 24168 GCTGGTCAACCCGCTGGAGCCCAAGCGCTTTTGGGGGCTCCGCCACACATCACCACA 24227
Qy 645 GGGCATCGTCTCTCAGCAGCAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 704
Db 24228 GGGCATCGTCTCTCAGCAGCAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 24287
Qy 705 AGGTAGCAAGTGGCGCCACTCGAACAGAGCAGACAAAGGGCCCTCTGTTGTTCCTCCCTCCC 764
Db 24288 AGGTAGCAAGTGGCGCCACTCGAACAGAGCAGACAAAGGGCCCTCTGTTGTTCCTCCCTCCC 24347
Qy 765 CATCTCCAGCAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTTGGAGGGAGGG 824
Db 24348 CATCTCCAGCAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTTGGAGGGAGGG 24407
Qy 825 CGCCCTTCAGAAAGCTGCAGCGCAGCTTTTGAAGGAGAGAGCTTCCTCCAGCCTGGCCCTA 884
Db 24408 CGCCCTTCAGAAAGCTGCAGCGCAGCTTTTGAAGGAGAGAGCTTCCTCCAGCCTGGCCCTA 24467
Qy 885 CGAGAGCGGCGCGCGCTGACGGAGCAGAGCTGGAGGGCCCGAGGCCCAAGGGCGCAA 944
Db 24468 CGAGAGCGGCGCGCGCTGACGGAGCAGAGCTGGAGGGCCCGAGGCCCAAGGGCGCAA 24527
Qy 945 CAAGCTCAAGCAGCCTCCGAGAGAGCCAGCGCGCCAGCAGCTCTGACACCTGCAGGT 1004
Db 24528 CAAGCTCAAGCAGCCTCCGAGAGAGCCAGCGCGCCAGCAGCTCTGACACCTGCAGGT 24587
Qy 1005 ACTCAGCTTCAGCAGGAGAGCGGCAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGGA 1064
Db 24588 ACTCAGCTTCAGCAGGAGAGCGGCAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGGA 24647
Qy 1065 GCAGAGCTGTGTGAGACCAAGCTCAGTCTCCTACGAGGGAGAGACACAGCTTCGGGCC 1124
Db 24648 GCAGAGCTGTGTGAGACCAAGCTCAGTCTCCTACGAGGGAGAGACACAGCTTCGGGCC 24707

```

QY 1125 CGCGCTGAGGAGACCCAGTGGAGGTGTGCC 1156
Db 24708 CGCGCTGAGGAGACCCAGTGGAGGTGTGCC 24739

RESULT 7
AF123653 Homo sapiens 9108 bp DNA linear PRI 07-APR-1999
LOCUS AF123653 Homo sapiens FEZ1 (FEZ1) gene, complete cds.
DEFINITION AF123653
ACCESSION AF123653
VERSION AF123653.1 GI:4572463
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9108)
AUTHORS Ishii.H., Baffa.R., Nunata.S.I., Murakumo.Y., Rattan.S., Inoue.H.,
Mori.M., Fidanza.V., Alder.H. and Croce.C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 9108)
AUTHORS Ishii.H., Baffa.R., Nunata.S.I., Murakumo.Y., Rattan.S., Inoue.H.,
Mori.M., Fidanza.V., Alder.H. and Croce.C.M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 2335 10th street, Philadelphia, PA 19107, USA
FEATURES
source
1. .9108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
<112..>5553
/gene="FEZ1"
join(<112..456,1707..2510,4912..5553)
/gene="FEZ1"
/product="FEZ1"
join(112..456,1707..2510,4912..5553)
/gene="FEZ1"
/codon_start=1
/product="FEZ1"
/db_xref="GI:4572464"
/translation="MGSVSLISGHSFKHCRASQYKLRKSHLKKLNRYSDGLLRP
GFSQDSHGKSSKMGSEDFYIKVQKARGSHHPDYATLSSDGLGGQAGVDFDPT
PPKLMPSNQLMEGSEKAVPTAFKPVLPKRSQALHSSPESASHQLHPAPDPKPEQ
ELKPLCSGALSDSGRNMSLPTSTSSSYOLDPLVTPVGPSTRFGGSAHNTQGV
LQDSNMSLKALSFSDGSKLGHKNKADKGPSCVRSPISTDECISIOLEOKLIEREGA
LQKQRFSEELASSLAYERPRRCRDELEGPEPKGNKLIKQASQSQRAQVHLIQ
VLQKQERQRQURQESLEKMDLLETKLRSYERKTSFGPALBETQWEVCQKSGEIS
LLKQOLKESEVENAKSEILGLKQAKDRTKLEGLERLTDQDLGALRTKGLLEVC
ENELQRKAEALREKRVNLEQLQELRAQALARDMDGPPPTFFPDVPAQLELELR
AELREERQHDQMSGFOHERLVNKEEKEKVIQYQKQLOQSVAMYQNRQRLXALQ
LARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT 2021 a 2628 c 2504 g 1952 t. 3 others
ORIGIN

Query Match 45.1%; Score 807.2; DB 9; Length 9108;
Best Local Similarity 99.6%; Pred. No. 7.5e-105;
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 345 GGGCTCCAGAGAGGTGAGTGGAGCCACAGCCTTCAAGCTGTGCTGCACAGGTTCAGG 404
Db 1706 GGGCTCCAGAGAGGTGAGTGGAGCCACAGCCTTCAAGCTGTGCTGCACAGGTTCAGG 1765

QY 405 AGCCATCTGACCTCTCCCGGAGAGTGCAGCCACCAGCTGCACCCCGCCCTCCAGA 464
Db 1766 AGCCATCTGACCTCTCCCGGAGAGTGCAGCCACCAGCTGCACCCCGCCCTCCAGA 1825

```


QY 825 CCGCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTGCTCCAGCTCGCCTA 884
 DB 84851 TGCACTACAGAGCTACAGCGCAGTTTCATGAGAAGAGTTGCTCTGGCCAAACCTT 84910
 QY 885 CGAGAGCGCGCGCGCTGCAGGAGCAGCTGCAGGCGCGGAGCCCAAGGCGGCAA 944
 DB 84911 TGAGAGCGCGCACCGGCGACTAGAGAGCACTGGAGTGCTTGGAGCCTAA-----GAG 84964
 QY 945 CAAGCTCAAGCAGGCTTCGCGAAGAGCAGCGCGCGCAGCAGGTCTCCACCTGCAGGT 1004
 DB 84965 CAAGCTGAAGCCACCTTCACAGAAGCAGCCAGCCACACAGCAGGTGCTCCAGCTGCAGGT 85024
 QY 1005 ACTGAGCTTCAGCAGGAGAGCGCGCAGCTCCGCGAGGAGCTTCAGAGCTCATGAAGA 1064
 DB 85025 GTTGCACTTCAGCAGGAGAAACCGCAGCTAAGCGAGGAGCTTAGAGAGCCTAAATGAAGA 85084
 QY 1065 CGAGACCTGTGGAGACCAAGCTCAGGTCCCTACGAGGAGGAGAGCAGCTTCGGGCC 1124
 DB 85085 CGAGACCTGTGGAGACCAAGCTCAGGTCCCTACGAGGAGGAGAGCAGCTTCGGGCC 85144
 QY 1125 CGCGCTGAGGAGACCCAGTGGGAGGTGTGCCA 1157
 DB 85145 TGCACTGGAGAGACCCCAATGGAGGTAAGACA 85177

RESULT 9
 AC099416/c
 LOCUS 263546 bp DNA linear HTG 23-MAY-2002
 DEFINITION Mus musculus chromosome UNK clone RP23-122M11, WORKING DRAFT
 SEQUENCE, 8 unordered pieces.
 AC099416
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Mus musculus.
 SOURCE Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 McPherson,J.D. and Waterston,R.H.
 The sequence of Mus musculus clone
 2 (bases 1 to 263546)
 McPherson,J.D. and Waterston,R.H.
 Direct Submission
 Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 263546)
 McPherson,J.D. and Waterston,R.H.
 Direct Submission
 Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 On May 23, 2002 this sequence version replaced gi:16924178.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@wustl.edu
 ----- Project Information -----
 Center project name: M_BA0122M11
 ----- Summary Statistics -----
 Sequencing vector: M13, 32x
 Sequencing vector: plasmid, 68x
 Chemistry: Dye-terminator Big Dye, 100% of reads
 Chemistry: Dye-terminator Big Dye, 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 278310 bases at least Q40
 Consensus quality: 282938 bases at least Q30
 Consensus quality: 286149 bases at least Q20
 Insert size: 250000; agarose-fp
 Insert size: 387570; sum-of-ctnigs
 Quality coverage: 25.74 in Q20 bases; agarose-fp

Quality coverage: 16.86 in Q20 bases; sum-of-ctnigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1204: contig of 1204 bp in length
 * 1205 1304: gap of unknown length
 * 1305 1524: contig of 1220 bp in length
 * 2525 2624: gap of unknown length
 * 2625 4136: contig of 1512 bp in length
 * 4137 4236: gap of unknown length
 * 4237 16434: contig of 12198 bp in length
 * 16435 16535: gap of unknown length
 * 32106: contig of 15572 bp in length
 * 32107 32206: gap of unknown length
 * 32207 140224: contig of 108018 bp in length
 * 140225 140325: gap of unknown length
 * 262693: contig of 122369 bp in length
 * 262694 262794: gap of unknown length
 * 262794 263546: contig of 753 bp in length.

FEATURES
 Location/Qualifiers
 1..263546
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP23-122M11"
 1..1204
 /note="assembly_name:Contig127"
 1305..2524
 /note="assembly_name:Contig157"
 2625..4136
 /note="assembly_name:Contig360"
 4237..16434
 /note="assembly_name:Contig396"
 16535..32106
 /note="assembly_name:Contig397"
 32207..140224
 /note="assembly_name:Contig398"
 140325..262693
 /note="assembly_name:Contig399"
 262794..263546
 /note="assembly_name:Contig154"
 710 others

BASE COUNT 76628 a 56149 c 55623 g 74436 t

ORIGIN
 Query Match 33.1%; Score 592.6; DB 2; Length 263546;
 Best Local Similarity 84.0%; Pred.No. 7.8e-75;
 Matches 683; Conservative 0; Mismatches 124; Indels 6; Gaps 1;
 QY 345 GGGTCGCGAAGGGTGCAGTGAGGCGCCACAGCCTTCAAGCTGTGCTGCCACGGTCAGG 404
 DB 250515 GAGTTACAGACAGGGGCGAGTAAAGCCCACTGCTTCAAGCCTGTGCTGCCACGGTCAGG 250456
 QY 405 AGCATCTCTGACTCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGA 464
 DB 250455 AGCATCTCTGACTCTCTTCACTGAGAGTACAGCCACCACTTCATCCATGCTCCAGA 250396
 QY 465 CAAGCCCAAGGAGCAGGCTGAAGCTTGCTGCTGCTGGGGCGCTGTCCAGACTCGG 524
 DB 250395 TAAGCCCAAGGAGCAGGAGCTGAAGCCAGGCTGTGCTCCGGGGCGCTGTCTGACTCAGG 250336
 QY 525 CCGGAATCTCATGTCTCAGCTGCGCCACACAGCAGCAGCAGCTTACAGCTGGACCC 584
 DB 250335 CCGGAATCTCATGTCTAGCTGCGCCAGCGATAGCACACACAGCAGCTTACAGCTGGATCC 250276
 QY 585 GCTGGTGCACACCGTGGGACCCCAAGCCCTTTTGGGGGCTCCGCCCAACATCACCA 644
 DB 250275 TCTGGTGCACCCCGTGGGGCTTACAGCCCGCTTTTGGGGGTTTCTGCTCACAACATCACCA 250216

```

QY 645 GGGCATCTCTCCAGGACAGCAATCATGAGCTGAAGCTCTGTCTCTCCGACGG 704
Db 250215 AGGCATCATCTTTCAGGACAGTAATATGATGAGCTGAAGCTCTGTCTCTGATGG 250156

QY 705 AGGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGCCCTCTGTGTCCGCTCCCC 764
Db 250155 CGGCAGCAAGCTGGCTCACCCAGGCAAGCAGATGAAGGGCCCTCTGTGTCTCC 250096

QY 765 CATCTCCAGGACGAGTGCAGCATCCAGGAGTGGAGCAGAAGCTGTTGGAGAGGAGG 824
Db 250095 ACTCTCCAGGATGAGTGCACCATCCAGGAGCTGGAGCAGAAGCTGCTGCAGAGGAG 250036

QY 825 CGCCCTCAGAGCTGCAGCGCAGCTTGGAGAGAGGAGCTGCTCCAGCTGGCCTA 884
Db 250035 TGCACTCAGAGCTTACAGCGCAGTTTCGATGAGAAGGAGTTGCTCTGCGCAACCTT 249976

QY 885 CGAGAGCGGCGCGCGCTGCAGGAGCAGCTGGAGGCGCGGAGCCCAAGCGGCA 944
Db 249975 TGAGAGCGGCGCACCGCGCACTAGAGAGCACTGGAGTGCCTGGAGCCTTA-----GAG 249922

QY 945 CAAGCTCAAGCAGGCTCCGAGAAGCCAGCGCGCAGCAGGTCTTGCACCTGCGAGGT 1004
Db 249921 CAAGCTGAAGCCACCTCACAGAAGCCAGCGCACACAGCAGGTGCTCCAGCTGCGAGT 249862

QY 1005 ACTGAGCTCAGCAGGAGAGCGCAGCTCCGCGAGGAGCTCGGAGGCTCATGAAGGA 1064
Db 249861 GTTGAGCTTGCAGAGGAGAAACGGCAGCTAAGGAGGAGCTAGAGAGCCTTAATGAAGGA 249802

QY 1065 GCAGGACCTCTGGAGACCAAGCTCAGGTCTTACAGAGGAGAGACAGCTTCGGGCC 1124
Db 249801 GCAGGACCTCTGGAGACCAAGCTCAGGTCTTACAGGCGGGAGAAACCACTTCGCCCC 249742

QY 1125 CGCGCTGGAGGAGACCCAGTGGAGGTGTGCCA 1157
Db 249741 TGCACCTGGAGGAGACCCCAATGGGAGGTAAAGCA 249709

RESULT 10
AF123654
LOCUS
DEFINITION Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
complete cds.
ACCESSION AF123654
VERSION AF123654.1 GI:4572465
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
JOURNAL 9919287
MEDLINE 10097140
PUBMED
REFERENCE 2 (bases 1 to 633)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
Direct Submission
Submitted (23-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 2335 10th street, Philadelphia, PA 19107, USA
FEATURES
Location/Qualifiers
1..633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="E264162"
1..633
/gene="FEZ1"

```

CDS

```

1..633
/gene="FEZ1"
/notes="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23835.1"
/db_xref="GI:4572466"
/translation="MGSVSSLISGHSFPHSKRASQYKLRKSSHLKLNRYSDGLLRF
GFQSDSGHGSKSMKMSKSEDFYIKVQSKARGSHHPDYTALSSDGLGQAGVDFDPSI
PPKMPFNSQLEMGSEKGAIVPTAFKPLFRSGAILHSSPESASHQHPAPDPKPEQ
ELKPLGCSGALSISGRNSSLRPSLEVDLEGADIPYEDIIATEI"
BASE COUNT 136 a 217 c 175 g 105 t
ORIGIN
Query Match 31.4%; Score 562.6; DB 9; Length 633;
Best Local Similarity 97.6%; Pred. No. 3.8e-70;
Matches 571; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 ATGGGCGAGCGTCACTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCGGGCT 60
Db 1 ATGGGCGAGCGTCACTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCGGGCT 60
QY 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAAGTCAACCGGTATTCGACGGG 120
Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAAGTCAACCGGTATTCGACGGG 120
QY 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGAGTCAAGTCCCAAAATGGGC 180
Db 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGAGTCAAGTCCCAAAATGGGC 180
QY 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 241 GATTACAGGCACTGTCCAGCGGGAATTTAGGGGGCCAGGCTGGGGTGGAGCTTTCACCG 300
Db 241 GATTACAGGCACTGTCCAGCGGGAATTTAGGGGGCCAGGCTGGGGTGGAGCTTTCACCG 300
QY 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAACTCAGTCTAGAAATGGGCTCCGAGAGG 360
Db 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAACTCAGTCTAGAAATGGGCTCCGAGAGG 360
QY 361 GCAGTGAAGGCCCCACAGAGCTTCAAGCCTGTGTGTCACAGCTCAGGAGGAGGAGGAGG 420
Db 361 GCAGTGAAGGCCCCACAGAGCTTCAAGCCTGTGTGTCACAGCTCAGGAGGAGGAGGAGG 420
QY 421 TCCCGGAGAGTGCAGCCACAGCTGACACCCCGCCCTCCAGACAGCCCAAGAGGAGGAG 480
Db 421 TCCCGGAGAGTGCAGCCACAGCTGACACCCCGCCCTCCAGACAGCCCAAGAGGAGGAGGAG 480
QY 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTCCAGACTCCGGCCGGAACCTCCATGTCC 540
Db 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTCCAGACTCCGGCCGGAACCTCCATGTCC 540
QY 541 AGCTGCTCCACACACAGCAGCAGCAGCTTACAGCTGGAGCCCG 585
Db 541 AGCTGCTCCACACACAGCAGCAGCAGCAGCTTGGAGGTTGACCTG 585

RESULT 11
AC108987/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-115K1, *** SEQUENCING IN PROGRESS
***, 78 unordered pieces.
ACCESSION AC108987
VERSION AC108987.3 GI:21737647
KEYWORDS HTG; HTGS PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 191210)

```


AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oesman, F.R., Allen, C., Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbatia, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Housi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lorado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Kapue, P., Martin, R., Martindale, A., Martinez, E., Masey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwu, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

REFERENCE

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 191210)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18846600.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPMW
Center clone name: CH230-115K1
----- Summary Statistics
Sequencing vector: Plasmid,
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115719 bases at least Q40
Consensus quality: 122885 bases at least Q30
Consensus quality: 129025 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 78 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1041: contig of 1041 bp in length
* 1042: gap of unknown length
* 1141: gap of unknown length
* 1142: contig of 1121 bp in length
* 2262: gap of unknown length
* 2363: contig of 1176 bp in length
* 3539: gap of unknown length
* 4775: contig of 1137 bp in length
* 4875: gap of unknown length
* 6050: contig of 1175 bp in length
* 6051: gap of unknown length
* 7660: contig of 1510 bp in length
* 7661: gap of unknown length
* 7761: contig of 1571 bp in length
* 9332: gap of unknown length
* 9432: contig of 1596 bp in length
* 11027: gap of unknown length
* 11028: contig of 1642 bp in length
* 11128: gap of unknown length
* 12869: gap of unknown length
* 12870: contig of 1223 bp in length
* 14092: gap of unknown length
* 14192: contig of 1036 bp in length
* 14193: gap of unknown length
* 15229: contig of 1381 bp in length
* 15329: gap of unknown length
* 16710: gap of unknown length
* 16809: contig of 1386 bp in length
* 18195: gap of unknown length
* 18295: contig of 1117 bp in length
* 18412: gap of unknown length
* 19512: contig of 1447 bp in length
* 20959: gap of unknown length
* 21059: contig of 1102 bp in length
* 2161: gap of unknown length
* 22162: contig of 1326 bp in length
* 22262: gap of unknown length
* 23587: contig of 1544 bp in length
* 25231: gap of unknown length
* 25331: contig of 1804 bp in length
* 27135: gap of unknown length
* 27235: contig of 1940 bp in length
* 29175: gap of unknown length
* 30846: contig of 1571 bp in length
* 30847: gap of unknown length
* 30946: contig of 1144 bp in length
* 32090: gap of unknown length
* 32091: contig of 1849 bp in length
* 32191: gap of unknown length
* 34039: contig of 1269 bp in length
* 34140: gap of unknown length
* 35508: contig of 1510 bp in length
* 35509: gap of unknown length
* 37018: contig of 1142 bp in length
* 37119: gap of unknown length
* 38260: contig of 1723 bp in length
* 38261: gap of unknown length
* 40083: contig of 1017 bp in length
* 40183: gap of unknown length
* 41200: contig of 2065 bp in length
* 41300: gap of unknown length
* 43365: contig of 1878 bp in length
* 43465: gap of unknown length
* 45343: contig of 1544 bp in length
* 45443: gap of unknown length
* 46987: contig of 2317 bp in length
* 47087: gap of unknown length
* 49404: contig of 1319 bp in length
* 49504: gap of unknown length
* 49505: contig of 1319 bp in length

```
* 50824 50923: gap of unknown length
* 50924 52451: contig of 1528 bp in length
* 52452 52551: gap of unknown length
* 52552 54441: contig of 1890 bp in length
* 54442 54541: gap of unknown length
* 54542 56442: contig of 1901 bp in length
* 56443 56542: gap of unknown length
* 56543 57749: contig of 1207 bp in length
* 57750 57850 60138: gap of unknown length
* 57850 60138: contig of 2289 bp in length
* 60139 60238: gap of unknown length
* 60239 62467: contig of 2229 bp in length
* 62468 62567: gap of unknown length
* 62568 64364: contig of 1797 bp in length
* 64365 64464: gap of unknown length
* 64465 66889: contig of 2425 bp in length
* 66890 66989: gap of unknown length
* 66990 68830: contig of 1841 bp in length
* 68831 68930: gap of unknown length
* 68931 71226: contig of 2296 bp in length
* 71227 71326: gap of unknown length
* 71327 73712: contig of 2386 bp in length
* 73713 73812: gap of unknown length
* 73813 76871: contig of 3059 bp in length
* 76872 76971: gap of unknown length
* 76972 79844: contig of 2873 bp in length
* 79845 79944: gap of unknown length
* 79945 81259: contig of 1315 bp in length
* 81260 81359: gap of unknown length
* 81360 83690: contig of 2331 bp in length
* 83691 83790: gap of unknown length
* 83791 86195: contig of 2405 bp in length
* 86196 86295: gap of unknown length
* 86296 88065: contig of 1770 bp in length
* 88066 88165: gap of unknown length
* 88166 90384: contig of 2219 bp in length
* 90385 90484: gap of unknown length
* 90485 91999: contig of 1515 bp in length
* 92000 92099: gap of unknown length
* 92100 94583: contig of 2484 bp in length
* 94584 94684 98297: contig of 3614 bp in length

Query Match 26.5%; Score 475.4; DB 2; Length 191210;
Best Local Similarity 83.4%; Pred. No. 2.8e-58;
Matches 561; Conservative 0; Mismatches 91; Indels 21; Gaps 1;

QY 1140 CCAGTGGGAGGTGTCCAGAGTCCAGGAGATCCCTCTCTGAAGCAGCAGCTGAAGGA 1199
DB 106039 CCTTGGCAGGTGTCCAGAGTCCAGGAGTCTCTCTCTGAAGCAGCAGCTGAAGGA 106000

QY 1200 GTCCAGACGGAGGTGAACGCCAGGCTAGCGAGATCTCTGGTCTCAAGGCACAGCTGAA 1259
DB 105999 ATCCCAATGGAGTGAATGCTTAAGCCAGTGCAGATCTCTAGCTTGAAGGCACAGCTGAA 105940

QY 1260 GGACACGGGGCAGCTGGAGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCT 1319
DB 105939 GGACACCGGGGCAAGCTTGGAGGATGGAGTTGAAGACACAGGACTTGGAGAGCGCCCT 105880

QY 1320 GCGCACCAAGGCGCTGGAGCTGGAGTCTGTGAGAAATGAGTGTGAGCGCAGCAGCTGAAGGA 1379
DB 105879 GCGCACCAAGGCGCTGAGCTGAGGCTGTGTGAGAAATGAGTGTGAGCGCAGCAGCTGAAGGA 105820

QY 1380 GCGGAGCTGTGCGGGAGAGAGGTGAACCTGCTGGAGCAGGAGCTGCGAGGCTCGGGGC 1439
DB 105819 GCGGAGCTGTGAGAGAGAGGTGAACCTTTTGAACAGGAGCTGCTGGAGCTCGGGGC 105760

QY 1440 CCAGGCGCGCTGGCGCCGCGCATGGGCC-----GCCCCACCTT 1478
DB 105759 CCAGGCGCGCTGCGATCGAGATGCGAGCACCTCGGGCCCTCAGGGATTTGGGCTGACTTT 105700

QY 1479 CCCCGAGGACGTCCCTGCGCTGACGCGGAGCTGGAGCGGCTGCGGGCGGAGCTCGGGGA 1538
```

```
Db 105699 CTCAGAGGACATCCCTGCTCTGCAGCGGGAGCTGGATCGGCTGCGGCTGAGCTGAAGGA 105640
QY 1539 GGAGCGCAAGGCCATGACAGATGCTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAA 1598
Db 105639 GGAGAGCAAGGCCATGACAGATGCTCTCGGGATTCAGCATGAGCGGTTGGTGTGGAA 105580
QY 1599 GGAGGAGAAGGAGAGGTGATTACAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCAT 1658
Db 105579 GGAAGAGAAGGAGAAAGTATCCAGTACCAAAACAGCTGCAACAGAGCTACTTGGCCAT 105520
QY 1659 GTACCAGCGGAACAGCGCCTGGAGAAAGCGCTGCAGCAGCTGCACGTGGGGACAGCGC 1718
Db 105519 GTACCAGCGCAACAGCGCTCTAGAGAAAGCCCTGCAGCAGCTGCGCCGTGGGGATATTGC 105460
QY 1719 CGGGAGCCCTTGGAGGTTGACCTTGGAAAGGCGCTGACATCCCTTACAGGACATCATAGC 1778
Db 105459 AGGGAGCCCTTTGAAATTGACCTGGAGGAGCTGACATCCCTTATGAGGATATCATAGC 105400
QY 1779 CACTGAGATCTGA 1791
Db 105399 CACTGAGATCTGA 105387
```

RESULT 12

AY029201

LOCUS

AY029201

DEFINITION

AY029201

ACCESSION

AY029201.1

VERSION

GI:15193289

KEYWORDS

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens.

REFERENCE

1 (bases 1 to 1935)

AUTHORS

Cabeza-Arvelaiz, Y., Thompson, T.C., Sepulveda, J.L. and Chinault, A.C.

TITLE

LAPSER 1 a novel candidate tumor suppressor gene on 10q24.3; cDNA

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1935)

AUTHORS

Cabeza-Arvelaiz, Y., Thompson, T.C., Sepulveda, J.L. and Chinault, A.C.

TITLE

Direct Submission

JOURNAL

Submitted (03-APR-2001) Molecular and Human Genetics, Baylor

FEATURES

Location/Qualifiers

1..1935

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

/map="10q24"

/tissue_type="normal human prostate"

1..1935

/gene="LAPSER1"

/note="putative tumor suppressor gene similar to LZTS1

(FEZ1)"

1..1935

/gene="LAPSER1"

/note="structural topology resembles bzip transcription

factors (CREB/ATF)"

/codon_start=1

/product="LAPSER1"

/protein_id="AAK31577.1"

/db_xref="GI:15193290"

/translation="MGSVSSLSISGRPCPGPAPPRHHPGPTFFROODGLLRGYEA

QEPICPAVPRKAVPTVSTFYINEDFTESPSPSDVEDAEQRAHNAHURGPPKL

IPVSGLEKWEKILIRPTAFKPVLPKPGAPSLPSFMGPRATGLSGQSLQTLFGG

PASSSSSSSSAADKPLAFSGWAGCPSTLSDSGRNSLSLPTSTGGAEPTTSSP

GGHLPSGSGRGAALPGARGVPTGSHSDSGRSSSKSTGSLGGRVAGLLGSGTRAS

PDSSCGERSFPFPPPPSDEALLHCIVLEKLRDREAELQOLRSLDSENEATWCQAYE

EROSHWOREAREALDCAQAQALQLOLQVFOLOQEKRLQDDFAOLLQREOQ

LERRCATLREORELGRLEETKWEVCQKGEISLLKQQLKESQAEILVQKSEIVALR

VALREARATLRVSEGRAGLQEAARELELEACQBLQRHROEAQLREKAGLDAAE

BASE COUNT
ORIGIN

355 a	619 c	668 g	293 t
-------	-------	-------	-------

AAGLRPPVPATADPELLAESDAKVORAAAAGVGSLSRAQVRLRVLEQRRRRRGEE
QRUSFEGERLAWQAEKQSVIRYKQLOHNYIQMYRRNRLEQEQLQQLSLEARELAD
LGLAEAPCICLEBITATEI"

Query Match	17.5%	Score 314;	DB 9;	Length 1935;
Best Local Similarity	59.6%	Pred. NO. 3.7e-35;		
Matches 617;	Conservative	0;	Mismatches 365;	Indels 54;
				Gaps 3;

Qy	796	CTGAGCAGAAAGCTGTTGGAGAGGAGGGCCGCTTCAGAAAGCTGACAGCGCAGCTTTGAG	855
Db	910	CTGAAAGAAAGCTCCGAGAGCCGGGAGGACAGAGCTTCAGCAGCTGCGGGACAGTCTCGAC	969
Qy	856	GAGAAAGAGCTTGCTCCAGCTCGCTTACGAGAGCGGCGCGCGCTGCAGGAGCAG	915
Db	970	GAGAAATGAGGCTACCATGTGCCAGCATACGAGAGCGGACGCGCACTTGGCAGCGAG	1029
Qy	916	CTGAGGGCCCGGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCGCTCGCAGAAAGGCCAG	975
Db	1030	CGTGAGGCCCTCGCAGAGGACTGTGCGGCCCAGGCACAGCGGC-----ACAG	1077
Qy	976	CGCCGCGACAGAGTCTTGCACTTCGACGTTCTGCAAGTTCGAGCTTCAGCAGGAGAAAGCGCAGCTC	1035
Db	1078	CGGGCCCAACAGCTGCTCGACTTCAGGTGTTCCAGCTGCAAGCAGGAGAAAGCGGCAATTG	1137
Qy	1036	CGGCAGAGCTCGAGAGCCTCATGAAGGACGAGCACTGCTGGAGACCAAGCTCAGGTCTC	1095
Db	1138	CAGACGACTTCGCACAGCTGCTCGAGAGCGCCAAAGCTTGGAGCGCGCTGCGCCACCC	1197
Qy	1096	TACGAGAGGAGAAAGACACAGCTTCGGGCCCGCGCTGGAGGAGACCCAGTGGAGGTGTGC	1155
Db	1198	TTGAGCGGGAGCAGCGGGAGCTCGGCCGAGGCTTGAGGAGACCAAGTGGAGGTGTGC	1257
Qy	1156	CAGAAGTCAAGCGAGATCTCTCTCTGAAGCAGCAGCTGAAGAGTCCCAGACGAGGAGT	1215
Db	1258	CAGAAATCAGCGAGATCTCCCTCTCTGAAGCAGCAGCTGAAAGAGTCTCAGGCAGAGCTG	1317
Qy	1216	AAGCCAAAGGCTACGAGATCTCTGGGTCTCAAGCACAGCTGAAGGAGACCGCGGGCAAG	1275
Db	1318	GTGCAGAAAGGCGCAGAGCTGGTGGCTCTGCGGTGGCGCTGCGGAGAGCCCGTGCTACG	1377
Qy	1276	CTGAGGGCCCTGGAGCTGAGGACCCAGGACCTTGAGGGCGCCCTGCGCACCAAGGGCCCTG	1335
Db	1378	CTCGGGTCAGTGAGGGCCGTGCGCGGGTCTACAGGAGCGGCCCGAGCTCGGAGCTG	1437
Qy	1336	GAGCTGAGAGTCTGTGAGATGAGCTGAGCGCAAGAAGAAACAGGCGGAGCTGCTCGCG	1395
Db	1438	GAGCTGAAGCTGTTCTCCAGGAGCTGCAGGCACACCGCCAGGAAGCTGAGCAGCTCGCG	1497
Qy	1396	GAGAAAGCTGGBGACATGTTGGATGTGAGGCGCGCGGACTCGCGGAGCCCCCTGTGCCACCT	1439
Db	1498	GAGAAAGCTGGBGACATGTTGGATGTGAGGCGCGCGGACTCGCGGAGCCCCCTGTGCCACCT	1557
Qy	1440	-----CCAGGCGCCCTGCGCCCGCACATGGGGCGCCGCCACCTTC	1479
Db	1558	GCCACCCCTGACCAATTCTCTTGGCAGAGTGTATGAGGCCAAAGTGCAGCGGGCAGCA	1617
Qy	1480	CCGAGACGTCCTGTCGCTGCAAGCGGAGCTGAGCGGCTGCGGGCCGAGCTGCGGGAG	1539
Db	1618	GCCGGGTTGGGGGACGTTGCGGGCCACAGTGGAGCGATTGCGGGTGGAGCTGACAGCGG	1677
Qy	1540	GAGCGGCAAGCCATGACACAGATGTCCTCGGGCTTCAGCATGAGCGGCTCGTGTGAAG	1599
Db	1678	GAGCGCGCGGGGTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGCGGCTGCGCTGCGAG	1737
Qy	1600	GAGGAGAAGGAGAGGTGATTTCAGTACAGAAAAGCTGCAGCAGAGCTACGTGGGCCATG	1659
Db	1738	GCAGAAAGGNCAGGTGATCCGCTACAGAAAGCAGCTGCAGCACACTACATCCAGATG	1797
Qy	1660	TACAGCGGAAACAGCGCTGGAGAAAGGCCCTGACAGCTGGCACTGGGGACAGCGCC	1719
Db	1798	TACCGGCGCAACCGGACGTAGACAGGAGCTGACAGAGCTGACAGCTGAGCTGAGGAGG	1857

Qy	1720	GGGGAGCCCTTGGAGGTGACCTGG-----AAGGGGCTGACATCCCTACGAGGAGATC	1773
Db	1958	CGGGAGCTCGCTGACCTACCTGGGCTTGGCCGAGCAGCCCCCTGCATCTCCCTGGAGGAGATC	1917
Qy	1774	ATAGCCACTGAGATCT	1789
Db	1918	ACTGCTACTGAGATCT	1933

RESULT 13
HSM805394
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFp761K171 (from clone DKFp761K171).
ACCESSION AL834338
VERSION AL834338.1 GI:21739969
KEYWORDS
SOURCE human.

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 2275)
TITLE	Ansozge.W., Wirkner.U., Mewes.H.W., Weil.B. and Wiemann,S.
JOURNAL	Direct Submission
COMMENT	Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761k1711) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/ .

FEATURES	source
----------	--------

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="10"
/clone="DKFzp761k1711"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..3275
/gene="DKFzp761k1711"
<1..1597>
/gene="DKFzp761k1711"
/notes="similarity to FEZ1 (Homo sapiens), different
splicing pattern"
/codon_start=2
/product="hypotheetical protein"
/protein_id="CAD39005.1"
/db_xref="GI:21739970"

```

polva signal
2236...2241
LADLGAEQACICEEITATEI"
GEEQRUSFGERIAWQAEKEQVIRYQKQJQHNYIQMYRRNRQLEQBELQQLSUELEARE
DEAAGUGRPPPPPPATADFPFLAESDQKVAQAAAGVGSLSAQOVLRLVLOLORRRR
ALRVALUREARATLRVSEGRARGLOAARALEEACSOELQRHQEASQLEKAGQL
REQLERCATLEREQLPRLTEETKWCQKSGEISLLKQQLKESQALVQSGSLV
AYEBRQHWQREARELQDCAQRAQRAQQLQLQVQLOQKQEQQLQDDFAQLQLE
RASPDSSSCGERSPPPPPPDEALLHCVLKGLRDREAELQQLRDSLDENEATWCQ
SSPGHLP SHGSGRGALPGARGVPTGPSHDSGRSSSSKTSGLSGRVAGGLGSGGT
FGPSSSSSSSSASDXPLAFSGWASGCPGTISLDSGRNSSLSPVTVTGCAGET
/translation) = "EKILRPTAFKVPKPRGAPSLPSFMGPFRATGSGSGSGLUTQ
/w_alpha1 = GA:217535970

polyA signal

polVA site	2257
/gene= "DKFZp161K1111"	

BASE COUNT	455 a	697 c	739 g	384 f
/gene="DKFZp61K1111"				

Query Match

17.5%: Score 314: DB 9: Length 2275:

Best Local Similarity 59.6%; Pred. No. 3.6e-35;		Matches 617; Conservative 0; Mismatches 365; Indels 54; Gaps 3;	
Qy	796	CTGGAGCAGAGCTGTTGGAGAGGAGGCGCCCTCCAGAGAGCTGCACGCGAGCTTTTGAG	855
Db	572	CTGGAAGAAAGCTCCGAGACCGGAGGAGAGCTTCAGCAGCTGCGGACAGTCTGGAC	631
Qy	856	GAGAAGAGCTTGCTTCAGCTGCTGCTACAGAGAGCGGCGCGCTGCGAGGAGCGAG	915
Db	632	GAGAAATGAGGCTACCTGTTGCGAGGCTACAGAGAGCGGCGGCTGCGAGGAGGAG	691
Qy	916	CTGGAGGCGCGGAGCCCAAGAGGCGGCAAGCTCAAGCAGGAGCTCGCAGAGAGCGAG	975
Db	592	CTGAGGCGCTTCAGAGGAGCTGTCGCGCCAGGCAAGCGGCGC-----ACAG	739
Qy	976	CGCGCGCAGCAGGCTTCGACCTGAGAGTATCTGAGCTTCAGCAGGAGAGAGCGGAGCTC	1035
Db	740	CGGCGCCCAACAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCGAGGAGAGCGGCAATTG	799
Qy	1036	CGGAGGAGCTCGAGGCTCATGAGGAGCAGGAGCTGCTGAGGAGCCAGCTCAGCTCAGCTCC	1095
Db	800	CAGGACGACTTCGACAGCTGCTGAGGAGCGGCAAGCAGCTGAGGAGCGGCTGCGCCACC	859
Qy	1096	TACGAGAGGAGAGAGCAGCTTCGCGCCCGCTGAGGAGAGCCAGTGGGAGGTGTGC	1155
Db	860	TTGAGCGGAGCAGCGGAGCTGCGCGGAGGCTTGAGGAGCAAGTGGGAGGTGTGC	919
Qy	1156	CAGAAGTCAGCGGAGATCTCCTCTGAAGCAGCAGCTGAAGAGTGTCCAGCAGGAGGTG	1215
Db	920	CAGAAATCAGCGGAGATCTCCTCTGAAGCAGCAGCTGAAAGAGTGTCCAGGAGAGCTG	979
Qy	1216	AACGCCAAGCTAGGAGATCTGCTGCTCAAGCAGCAGCTGAAGGACACGCGGGGCAAG	1275
Db	980	GTGAGAGGAGCAGCAGCTGCTGCTGCTGCGGAGTGTGCGGAGGCGGCTGCTGCTACG	1039
Qy	1276	CTGGAGGCGCTGAGCTGAGGACCCAGGAGCTGAGGCGGCGCTGCGGAGCAGGAGCTG	1335
Db	1040	CTCGGGTCTGAGGCGGCTGCGGAGGCTTACAGGAGCGGCGGAGCTGCGGAGCTG	1099
Qy	1336	GAGCTGAGGCTGTGAGAGTGTGAGCTGAGCGGAGAGAGAGAGAGAGGAGGAGCTGCTGCGG	1395
Db	1100	GAGCTGGAAGCTGTTCAGGAGCTGCGGAGGAGCTGAGGAGAGCTGAGCAGCTGCGG	1159
Qy	1396	GAGAGGTGACCTGCTGAGCAGGAGCTGAGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT	1439
Db	1160	GAGAAAGTGGGAGTGGATGCTGAGGCGGCGGAGCTGCGGAGGAGCCCTGTGCGACCT	1219
Qy	1440	-----CCAGGCGCGCTGCGCGGAGCAGCTGCGGAGGAGGCGGCGGCGGAGCTTCC	1479
Db	1220	GCCACCGTGACCATCTCTCTGCGAGAGTGTGAGGCGGAGGAGGAGGAGGAGGAGGAG	1279
Qy	1480	CCGAGGAGCTGCTGCTGCGGAGGAGCTGAGGAGGAGCTGCGGAGGAGGAGGAGGAGGAG	1539
Db	1280	GCGGGGTTGGGGGAGCTGCGGAGGAGCTGAGGAGGAGCTGCGGAGGAGGAGGAGGAGGAG	1339
Qy	1540	GAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1599
Db	1340	GAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1399
Qy	1600	GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1659
Db	1400	CGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1459
Qy	1660	TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1719
Db	1460	TACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1519
Qy	1720	GGGAGGCGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1773
Db	1520	CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1579
Qy	1774	ATAGCCACTGAGATCT 1789	

Db 1580 ACTGCTACTGAGATCT 1595

RESULT 14

BC006212

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC006212 2766 bp mRNA linear PRI 12-JUL-2001

Homo sapiens, clone MGC:2586 IMAGE:3161855, mRNA, complete cds.

BC006212 1 GI:13623228

MGC.

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2766)

Strausberg, R.

Direct Submission

Submitted (09-APR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 5 Row: k Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, GenomeScan gene prediction.

Location/Qualifiers

1..2766

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:2586 IMAGE:3161855"

/tissue_type="Brain, neuroblastoma"

/clone_lib="NIH_MGC_19"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1138..2088

/codon_start=1

/product="Unknown (protein for MGC:2586)"

/protein_id="AAH06212.1"

/db_xref="GI:13623229"

/translation="MCQAYEERHWRERAREALDCAQAQRAQQLQLQVFLQ QEKROLQDPAQLLOERELERRCATLREQLGRLPRLETKWEVCCKGEISLLKQ QKESQALVQKSELVALRVALREARATLVSEGRAGRLQEARAELEACSOEL QRHQEAREQUREKQGLDAEAGRPVFPATADPPLAESEDAKQVRAAAGVGGSL RAQVRLRVELQRRERREGRDQSFEGELAWQAEQVIRYQKQLOHNYIQMYRRNR ALBQELQLSLEARELADLGLAEQAPCICLEITATEI"

BASE COUNT 545 a 873 c 883 g 465 t

ORIGIN

Query Match

Best Local Similarity

Matches 617; Conservative

17.5%; Score 314; DB 9; Length 2766;

59.6%; Pred. No. 3.5e-35;

0; Mismatches 365; Indels 54; Gaps 3;

```
Qy 796 CTGAGCAGAAAGCTTTGGAGAGGAGGCGCCCTCCAGAAAGCTGCAGCGAGCTTTGAG 855
Db |||||
1063 CTGGAAGGAAAGCTCCGAGACCGGAGCAGAGCTTACAGAGCTGCGGACAGTCTGGAC 1122
Qy 856 GAGAGGAGCTTGCTCCAGCTGCTTACAGAGAGCGCGCGCTGCAGGACGAG 915
Db |||||
1123 GAGATGAGGCTACCATGTGCGAGCATACGAGGAGCGGAGCGCACTGGCAGGAG 1182
Qy 916 CTGAGGCGCCGAGACCCAAAGGCGGCAACAGCTCAAGCAGCGGCTTCGACAGAGCCAG 975
Db |||||
1183 CGTGAGGCCCTCGAGAGAGACTGTGCGGCCCGAGCACAGCGGC-----ACAG 1230
Qy 976 CGCGCGAGAGCTCTCTCACTGAGAGTACTGAGCTTACAGAGGAGAGCGGAGCTC 1035
Db |||||
1231 CGGCGCCAAACAGCTGCTGAGCTGAGGTGTTCCAGCTGCAGCAGGAGGCGGAATTG 1290
Qy 1036 CGGAGGAGCTCGAGAGCTCATGAAGCAGCAGGACCTGCTGGAGACCAAGCTCAGGTCC 1095
Db |||||
1291 CAGACGACTTGCACAGCTGCTGAGGAGCGCGACAGCTGGAGCGGCTGCGCCACC 1350
Qy 1096 TACGAGGAGGAGAACAGCAGCTTCGGCCCCCGCTGGAGGAGACCCAGTGGAGGTGTC 1155
Db |||||
1351 TTGAGCGGAGCAGCGGAGCTCGGCGCGAGGCTTGAGGAGACCAAGTGGAGGTGTC 1410
Qy 1156 CAGAGTCAAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGGTCCAGACGAGGTG 1215
Db |||||
1411 CAGAAATCAGCGAGATCTCCCTGCTGAAGCAGCAGCTGAAAGAGTCTCAGCGCAGAGCTG 1470
Qy 1216 AACGCCAAGCTACGAGATCTTGGTCTCAAGGCACAGCTGAAGGACACCGGGGCAAG 1275
Db |||||
1471 GTGCAGAAAGGCGAGAGCTGGTGGCTCTGCGGTGCGCTGCGGGAGGCGGCTGTACTG 1530
Qy 1276 CTGAGGCGCTTGGAGCTCAGGACACCGAGCTGAGGCGCGCTCGCACCAAGGCGCTG 1335
Db |||||
1531 CTGCGGTCTAGTAGGCGCTGCGGGGCTCTACAGAGGCGCGCGAGCTCGGGAGCTG 1590
Qy 1336 GAGCTGAGGTCTGTGAGAAATGAGCTGAGCGCGAGAGAAAGAACGAGGCGAGCTGTCGG 1395
Db |||||
1591 GAGCTGGAAGCTGTGTCAGGAGCTGCGAGGACACCGCCAGGAAGCTGAGCAGCTGGCG 1650
Qy 1396 GAGAGGTGACCTCTGTGAGCAGAGCTGAGGAGCTGCGGCG----- 1439
Db |||||
1651 GAGAAAGTGGGAGTTGGATGCTGAGGCGCGCGAGCTCCGGAGCGCCCTGTGCCACT 1710
Qy 1440 -----CCAGCGCCCGCTGCGCGCGAGATGATGAGGCGCAAGTGACGCGGCGAGCA 1770
Db |||||
1711 GCCACCGTGACCATTCCTCTTGGCAGAGTATGAGGCGCAAGTGACGCGGCGAGCA 1770
Qy 1480 CCGAGGAGCTCCCTGCTGAGCGGAGCTGAGCGGCTGCGGCGCGAGCTGCGGAG 1539
Db |||||
1771 CCGCGGTTGGGGGAGCTTGGCGGCCAGGTTGAGCGGATTTGCGGGTGGAGCTGCGCGG 1830
Qy 1540 GAGCGGCAAGCCATGACAGATGCTCTGCGGCTTCCAGCATGAGCGCTGTTGGAG 1599
Db |||||
1831 GAGCGCGCGGGGTGAGGAGCGCGGAGCTTTGAGGGGAGCGGCTGCGCTGGCAG 1890
Qy 1600 CAGGAGAGGAGAAAGGTGATTGATCCAGAAAACAGCTGCAGCAGAGCTAGTGGCCATG 1659
Db |||||
1891 CGAGAGAGGAGAGGTGATCGGTACAGAGAGAGCTGAGACCAACTATCATCAGATG 1950
Qy 1660 TACCAGCGGAACCAAGCGCTGAGAGAGCCCTGAGCAGAGCTGGCAGCTGGGGAGCAGCGCC 1719
Db |||||
1951 TACCGGCGCAACCGCAGCTAGAGCAGAGCTGAGCAGAGCTCAGCTGAGCTGGAGGCC 2010
Qy 1720 GGGAGCGCTTGGAGGTGACTGG-----AAGGGGCTGACATCCCTTACGAGACATC 1773
Db |||||
2011 CGGAGCTGCTGACCTGGGCTGCGCGAGCAGGCGCCCTGCTCATCTGCTTGGAGGATC 2070
Qy 1774 ATAGCCACTGAGATCT 1789
Db |||||
2071 ACTGCTACTGAGATCT 2086
```

```
RESULT 15
LOCUS AB011124
DEFINITION Homo sapiens mRNA for KIAA0552 protein, complete cds.
ACCESSION AB011124
VERSION AB011124.1 GI:3043627
KEYWORDS KIAA0552 protein.
SOURCE Homo sapiens male brain cDNA to mRNA, clone lib:pBluescriptII SK
plus clone:HH0869.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro
JOURNAL DNA Res. 5 (1), 31-39 (1998)
MEDLINE 98290545
REFERENCE 2 (bases 1 to 5257)
AUTHORS Ohara,O., Nagase,T. and Ishikawa,K.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
DNA Technology; 1532-3 Yana, Kisearazu, Chiba 292-0812, Japan
(E-mail:cdmainfo@kazusa.or.jp. Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
Location/Qualifiers
source 1..5257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HH0869"
/sex="male"
/tissue type="brain"
/clone_lib="pBluescriptII SK plus"
gene 1..5257
/gene="KIAA0552"
CDS 1409..3430
/gene="KIAA0552"
/codon_start=1
/product="KIAA0552 protein"
/protein_id="BAA25478.1"
/db_xref="GI:3043628"
/translat="MAKLETPVRADPDRPDLAFAPRPSLGGPDPRLAMGSGV
AHQEPAMKVGTRTGGGSGGSGRGRPGYPSGDKGLANSLVNGE
LRGSDHTDVCNVVSGSSSGSGSKAPPOYREPSHPKLLATSKLDOCSPLVR
PSAFKPVVPKNFHSQMNLCPQNTGTEGQGLGGLDKSRTMTAGSGSLSD
SGRNLSTLSLTSYSSYSHLAPUSASTSHNRIGTASYGSGSGSGSGGYDGLTS
DSGRASKSGSSSMRPHLGSGGGGLPFAACSPSPSALIQLERLWEKEQ
VAALRRLEQSEAQAQVLEERQKAWERELAEFGSGKLOQVRAQRAQGLQLQ
VLRLQDKQLQBEAARLMRQBELEKVAQKEADFLPRIETKWEVCKAGEIS
LLKQOLKSDADYSOKLSEIVGLRSQLEGRASLREKEQLSLRDSFSSKQASLELG
EGELPAACLXPALTVPDPAEPDQALATCDEAKMRRAQAGVAAASLVSVDGEAEG
ESGTRALRRVGLQAEALAAARRARERQGSFAEERRVLEKEKVEIKYQKOLQSYV
EMTORNQLRRLRERGAAGGATPTPHQGEKKAWTPPSRLRIESTEI"
BASE COUNT 997 a 1618 c 1673 g 969 t
ORIGIN
Query Match 17.5%; Score 314; DB 9; Length 5257;
Best Local Similarity 58.2%; Pred. No. 3,1e-35;
Matches 619; Conservative 0; Mismatches 400; Indels 45; Gaps 2;
Qy 710 GCACGTGGGCCACTCGAACACAGGACAGAGGCCCTCGTGTCGCTCCGCCCTGCTCACCGC 2351
Db |||||
2292 GCCACCTGGGCTCTGGGAGGGCGGAGGTGAGGCCCTGCGCTTCGCGGCTGCTCACCGC 829
Qy 770 CCACGAGCAGGTGCAGCATCCAGGAGCTGGAGCAGAGAGCTGTTGGAGAGGAGGCGGCC 2411
Db |||||
2352 CCTCCCCAGTGCACTCATCCAGGAGCTGGAGGAGCGGCTGTTGGAGAGGAGGAGGAGG 889
Qy 830 TCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCGCTGCGCTACGAGG 2471
Db |||||
2412 TGGCAGCTCTGCGCGCGAGCGCTGAGCAGAGCGAGCGGCGCTGTGGCCCGAGGTACTGGAGG
```

```
Qy 890 AGCGCGCGCGCTGACAGGACGAGCTGGAGGCGCGGAGCCCAAGCGGCAACAGC 949
Db |||||
Qy 2472 AGCGGCAAGAGCGTGGAGCGGAGCTGGCCGAGCTGCGGACGGGTGCAGCGGAGC 2531
Db |||||
Qy 950 TCAAGCAGGCGCTGCGACAGAGCCAGCGCGCGAGCAGGTCTTCACCTGCAGGTACTGC 1009
Db |||||
Qy 2532 TACAGCAGTGGCGCGCGCTGCCAGCGCGCCAGCAGGCGCTACAGCTGCAGGTGTC 2591
Db |||||
Qy 1010 AGCTTACGAGGAGAGCGGCGAGCTCGGCGAGAGCTGCGAGAGCTCATGAGAGCAGG 1069
Db |||||
Qy 2592 GGCTGCACAGGACAAAGAGAGCTGCAGGAGAGGCGCGCGCTGATGCGGAGCGGG 2651
Db |||||
Qy 1070 ACCTGCTGGAGACCAAGCTCAGTTCCTACGAGAGGAGAGACAGCTTGGCCCCCGGC 1129
Db |||||
Qy 2652 AAGAGCTGGAGGACAAGGTGGCGCTGCGCAGAGGAGCAGCGGAGCTTCTGCCCCGGA 2711
Db |||||
Qy 1130 TGGAGGAGACCCAGTGGGAGGTGTGCCAGAACTCAGGCGAGATCTCCTCTGAGCAGC 1189
Db |||||
Qy 2712 TAGAGGAACTAAGTGGGAGGTGTGCCAGAGGCTGGCGAGATCTCCCTCTGAGCAGC 2771
Db |||||
Qy 1190 AGCTGAGGAGTCCAGACGAGGAGTGAACGCAAGGCTAGCGAGATCTGSGTCTCAAGG 1249
Db |||||
Qy 2772 AGCTGAAGGACTCGCAGCGGATGTGTGCGAAGTTGAGTGAGATCGTGGGACTGGGCT 2831
Db |||||
Qy 1250 CACAGCTGAAGGACACCGGGGCAAGCTGGAGGCGCTGAGAGCTGAGGACCCAGGACCTGG 1309
Db |||||
Qy 2832 CGCAGCTCGGGAGGCGCGGCTTCGCTGCGGGAGAGGAGGAGCAGCTGCTCAGCCCTGC 2891
Db |||||
Qy 1310 AGGCGCGCTCGCACCAAGGCGCTGGAGCTGGAGGTGTGAGAAATGAGCTGAGGCGCA 1369
Db |||||
Qy 2892 GGGACTCTTCAGCAGCAGCAGGCGGAGCTGGAGCTGGGCGAGGCGAGCTGCTGCCG 2951
Db |||||
Qy 1370 AGAAGAACGAGCGGAGTGTGCGGGAGAGAGGTGAACCTCTCTGAGCAGGAGCT---- 1424
Db |||||
Qy 2952 CTGCTCAAGCGCGCGCTGACCCCGCTGAGACCGCGGCGAGCCACAGGATCTCTGGCCA 3011
Db |||||
Qy 1425 -----GCAGGAGCTCGGGGCGCGGCGCGCTGAGGCGCGCTGAGGCGCGGCTGCGCTCCT 3071
Db |||||
Qy 1469 CGCCACCTTCCCGAGGAGC-----TCCCTGCCCTGCAGC 1504
Db |||||
Qy 3072 TGGTTCCGTGACGCGGAGCGGAGGCTGCGGGAGAGCGGAGCGGCGCGCTGCGGC 3131
Db |||||
Qy 1505 GGGAGCTGGAGCGGCTGCGGCGGAGCTGCGGGAGGAGCGGCAAGGCCATGACAGATGT 1564
Db |||||
Qy 3132 GGGAGGTGGGCGGCTGACAGCGGAGCTGCGGCTGAGCGCGGCGCGGAGCGCGGAGC 3191
Db |||||
Qy 1565 CCTGGGCTTCCAGCATGAGCGGCTCGTGTGAGAGGAGGAGAGAGGATTCAGT 1624
Db |||||
Qy 3192 GTGCCAGCTTCCCGAGGAGCGCGCTGTGCTGGAGAGAGGAGAGGATTCAGT 3251
Db |||||
Qy 1625 ACCAGAAACAGTGCAGCAGAGCTACGTGGCCATGTACAGCGGAACCCAGCGCTTGAGA 1684
Db |||||
Qy 3252 ACCAGAACAGCTGAGCTGAGCTACGTGGAGATGTACAGGCGAACCCAGCAGCTGGAGC 3311
Db |||||
Qy 1685 AGGCCCTGAGCAGCTGGCAGCTGGGAGCAGCGCGGGAGGCC 1728
Db |||||
Qy 3312 GCAGGCTGCGGAGCGCGGCGCGGAGGCGGTGCAAGCACGCC 3355
Db |||||
```

Search completed: June 15, 2003, 02:18:51
Job time : 4670.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:38:41 ; Search time 412.345 Seconds
(without alignments)
9781.434 Million cell updates/sec

Title: US-09-513-888C-3
Perfect score: 1791
Sequence: 1 atggcgagcgtcagtagcct.....tcatagcactgagatctga 1791

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1791	100.0	1791	21 AAA64509	CDNA sequence enco
2	1791	100.0	5492	21 AAA64508	CDNA sequence of c
3	1643	91.7	1722	21 AAA64515	Nucleotide sequenc
4	1581.4	88.3	1692	21 AAA64514	Nucleotide sequenc
5	1427	79.7	1614	21 AAA64512	Nucleotide sequenc
6	1372.4	76.6	1512	21 AAA64513	Nucleotide sequenc
7	808.8	45.2	2411	22 AAS30637	DNA encoding novel
8	808.8	45.2	2411	22 AAS28699	Genomic sequence #
9	807.2	45.1	9048	21 AAA64507	Nucleotide sequenc

10	562.6	31.4	633	21	AAA64511	Nucleotide sequenc
11	560	31.3	560	22	AAS30595	DNA encoding novel
12	560	31.3	560	22	AAS28127	Novel cDNA encodin
13	269.8	15.1	1623	24	ABO54361	Human ovarian anti
14	258.2	14.4	1617	21	AAF15625	Human prostate can
15	249	13.9	404	21	AAA64510	Nucleotide sequenc
16	148.4	8.3	657	22	AAF22599	Human breast cance
17	146	8.2	1926	21	AAA50254	Epstein Barr virus
18	146	8.2	1926	22	AAF82902	EBV tethering prot
19	146	8.2	2580	21	AAA75454	Nucleotide sequenc
20	146	8.2	2580	24	AAI64275	Epstein-Barr virus
21	146	8.2	5452	20	AAK50923	Anti-sense strand
22	146	8.2	8705	20	AAZ23778	Vector pshuttle DN
23	146	8.2	9600	19	AAV21683	Vector plasmid pCM
24	146	8.2	10380	20	AAZ22248	Nucleotide sequenc
25	146	8.2	10596	14	AAQ51731	Plasmid pCisEBON f
26	146	8.2	10596	17	AAAT40348	Plasmid pCisEBON f
27	146	8.2	10596	20	AAK15650	Nucleotide sequenc
28	146	8.2	16080	21	AAK59553	DNA clone pCEK Cl.
29	144.2	8.1	799	19	AAV55831	Nucleotide sequenc
30	138.4	7.7	3489	21	AAA30290	Kaposi's sarcoma-a
31	138.4	7.7	3489	22	AAF82901	Nucleotide sequenc
32	138.4	7.7	3489	24	ABA93487	Kaposi's sarcoma-a
33	138.4	7.7	32207	20	AAV73805	KSHV LUR DNA (nucl
34	138.4	7.7	137507	19	AAV19941	KSHV long unique c
35	134	7.5	1925	20	AAK50924	Epstein Barr Virus
36	131.4	7.3	795	19	AAV55830	FLGA insert stabl
37	131.2	7.3	2108	24	ABL67774	Oesophagus cancer
38	125.2	7.0	9551	20	AAZ22301	cDNA encoding a hu
39	118.2	6.6	2744	16	AAQ98470	MISPI-containing p
40	118	6.6	1588	22	AAH26304	Spider silk protei
41	117.4	6.6	2004	18	AAAT85356	Nephila clavipes s
42	117	6.5	2338	12	AAQ14183	N.clavipes draglin
43	114	6.4	561	24	ABK62287	Rat sequence diffe
44	113.8	6.4	2338	19	AAV23249	Nephila clavipes s
45	113.8	6.4	2338	21	AAZ38195	N. Clavipes spider

ALIGNMENTS

RESULT 1	
AAA64509	
ID	AAA64509 standard; cDNA; 1791 BP.
AC	AAA64509;
XX	
DT	02-JAN-2001 (first entry)
XX	
DE	CDNA sequence encoding a human FEZ1 polypeptide.
XX	
KW	Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW	tumour proliferation; tubulin; microtubule; protein BFL-gamma;
KW	tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW	cell growth; cell shape; cell rigidity; cell motility; DNA replication;
XX	tumorigenesis; tumour survival; metastasis; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1791
FT	/*tag= a
FT	/product= "FEZ1"
XX	
PN	WO2000050565-A2.
XX	
PD	31-AUG-2000.
XX	
PF	25-FEB-2000; 2000WO-US04950.
XX	
PR	25-FEB-1999; 99US-0121537.
XX	
PA	(UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Ishii H;
XX PI
XX WPI; 2000-558396/51.
DR P-PSDB; AAB08715.
XX
XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
XX Claim 7; Fig 5I; 255pp; English.
XX
XX The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a
CC tumour suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC E1-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;

Query Match 100.0%; Score 1791; DB 21; Length 1791;
Best Local Similarity 100.0%; Pred. No. 7,48-288;
Matches 1791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGACGGTCTAGTCTATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGCT 60
DB |||||
DB 1 ATGGGACGGTCTAGTCTATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGCT 60
QY 61 TCGCAGTACAGCTCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGGG 120
DB |||||
DB 61 TCGCAGTACAGCTCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGGG 120
QY 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGGTACCGGAAGTCCAGCTCCAAATGGGC 180
DB |||||
DB 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGGTACCGGAAGTCCAGCTCCAAATGGGC 180
QY 181 AAGACGGAAGCTTCTTCTACATCAAGTTCAGCCAGAAAGCCGGGCTCCCATCACCCTA 240
DB |||||
DB 181 AAGACGGAAGCTTCTTCTACATCAAGTTCAGCCAGAAAGCCGGGCTCCCATCACCCTA 240
QY 241 GATTACAGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
DB |||||
DB 241 GATTACAGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
QY 301 TCACACCCCAAGCTATGCCCTTCTCAATCAGTCTAGAAATGGGCTCCGAGAGGGT 360
DB |||||
DB 301 TCACACCCCAAGCTATGCCCTTCTCAATCAGTCTAGAAATGGGCTCCGAGAGGGT 360
QY 361 GCAGTGAGGCCACAGCTTCAAGCTGTGCTGCCACAGCTCAGGAGCCATCTGCACTCC 420
DB |||||
DB 361 GCAGTGAGGCCACAGCTTCAAGCTGTGCTGCCACAGCTCAGGAGCCATCTGCACTCC 420
QY 421 TCCCGGAGAGTGCCAGGCCACAGCTGCAACCCCGCCCTCCAGACAAAGCCCAAGAGCAG 480
DB |||||
DB 421 TCCCGGAGAGTGCCAGGCCACAGCTGCAACCCCGCCCTCCAGACAAAGCCCAAGAGCAG 480
QY 481 GAGCTGAAGCTGGCTGTGCTCTGGGGGCTGTGAGACTCCGGCCGGAACCTCCATGTCC 540
DB |||||
DB 481 GAGCTGAAGCTGGCTGTGCTCTGGGGGCTGTGAGACTCCGGCCGGAACCTCCATGTCC 540

QY 541 AGCCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCCCGCTGGTCCACACCCGTG 600
DB |||||
DB 541 AGCCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCCCGCTGGTCCACACCCGTG 600
QY 601 GGACCCCAAGCCGTTTGGGGCTCCGCCACAACATCACCCAGGGGATCGTCTCTCAG 660
DB |||||
DB 601 GGACCCCAAGCCGTTTGGGGCTCCGCCACAACATCACCCAGGGGATCGTCTCTCAG 660
QY 661 GACAGCAACATGATGAGCCTGAGGCTCTGCTCTTCTCCGACGAGGTAGCAAGCTGGC 720
DB |||||
DB 661 GACAGCAACATGATGAGCCTGAGGCTCTGCTCTTCTCCGACGAGGTAGCAAGCTGGC 720
QY 721 CACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGCTCCGCTCCCCATCTCCACGACGAG 780
DB |||||
DB 721 CACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGCTCCGCTCCCCATCTCCACGACGAG 780
QY 781 TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGTGAGAGGGAGGGCGCCCTCCAGAAGCTG 840
DB |||||
DB 781 TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGTGAGAGGGAGGGCGCCCTCCAGAAGCTG 840
QY 841 CAGCGCAGCTTGGAGAGAGGAGCTTGCCTCCAGCCTGGCCTACGAGGAGGAGCGCCGG 900
DB |||||
DB 841 CAGCGCAGCTTGGAGAGAGGAGCTTGCCTCCAGCCTGGCCTACGAGGAGGAGCGCCGG 900
QY 901 CGCTGCAGGAGCAGCTGGAGGGCCCGAGCCCAAGGCGGCAACAGCTCAAGCAGGCC 960
DB |||||
DB 901 CGCTGCAGGAGCAGCTGGAGGGCCCGAGCCCAAGGCGGCAACAGCTCAAGCAGGCC 960
QY 961 TCGCAGAAAGCCAGCGCGGCGCAGAGTCTGTCACCTGCAAGTCTGAGCTTTCAGCAG 1020
DB |||||
DB 961 TCGCAGAAAGCCAGCGCGGCGCAGAGTCTGTCACCTGCAAGTCTGAGCTTTCAGCAG 1020
QY 1021 GAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGGAGGAGCTCTCTGAG 1080
DB |||||
DB 1021 GAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGGAGGAGCTCTCTGAG 1080
QY 1081 ACCAAGCTCAGGTCCTACGAGAGGAGAGAACACAGCTTCGGCCCGCGCTCGAGGAGACC 1140
DB |||||
DB 1081 ACCAAGCTCAGGTCCTACGAGAGGAGAGAACACAGCTTCGGCCCGCGCTCGAGGAGACC 1140
QY 1141 CAGTGGAGGTGTCCAGAAAGTCAAGGAGATCTCTCTTGAAGAGCAGCTGAAAGAG 1200
DB |||||
DB 1141 CAGTGGAGGTGTCCAGAAAGTCAAGGAGATCTCTCTTGAAGAGCAGCTGAAAGAG 1200
QY 1201 TCCAGAGCGAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAGGGCAGCTGAAG 1260
DB |||||
DB 1201 TCCAGAGCGAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAGGGCAGCTGAAG 1260
QY 1261 GACACGGGGCAAGCTGGAGGGCTGGAGCTGAGGAGCCACAGGACCTGGAGGGCGCCCTG 1320
DB |||||
DB 1261 GACACGGGGCAAGCTGGAGGGCTGGAGCTGAGGAGCCACAGGACCTGGAGGGCGCCCTG 1320
QY 1321 CGCACCAAGGGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGGCGCAAGAACAGAG 1380
DB |||||
DB 1321 CGCACCAAGGGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGGCGCAAGAACAGAG 1380
QY 1381 GCGGAGCTGTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGCGAGGCTCGGGGCC 1440
DB |||||
DB 1381 GCGGAGCTGTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGCGAGGCTCGGGGCC 1440
QY 1441 CAGCGCCGCTGGCCCGCGCATATGGGGCCGCCACCTTCCCGAGGAGCTGCTGCCCCCTG 1500
DB |||||
DB 1441 CAGCGCCGCTGGCCCGCGCATATGGGGCCGCCACCTTCCCGAGGAGCTGCTGCCCCCTG 1500
QY 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCAAGGCATGACAG 1560
DB |||||
DB 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCAAGGCATGACAG 1560
QY 1561 ATGTCTCTGGGCTTCCAGCATGAGCGCTCTGTGTGAGAGGAGGAGAGGAAAGTGTATT 1620
DB |||||
DB 1561 ATGTCTCTGGGCTTCCAGCATGAGCGCTCTGTGTGAGAGGAGGAGAGGAAAGTGTATT 1620
QY 1621 CAGTACCAGAAACAGCTGCGAGAGCTACGTGGCCATGTACCAGCGGAAACAGCGCCTG 1680

Db 1621 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTACCGGAAACACGCGCCTG 1680
Qy 1681 GAGAAGGCCCTGTCAGCAGCTGGCAGCTGGGACACGCGGGGAGCCCTTTGGAGTTGAC 1740
Db 1681 GAGAAGGCCCTGTCAGCAGCTGGCAGCTGGGACACGCGGGGAGCCCTTTGGAGTTGAC 1740
Qy 1741 CTGGAAGGGGCTGATCATCCCTACGAGACATCATAGGACATGAGATCTGA 1791
Db 1741 CTGGAAGGGGCTGATCATCCCTACGAGACATCATAGGACATGAGATCTGA 1791

RESULT 2

ID AAA64508 standard; cDNA; 5492 BP.

AC AAA64508;

DT 02-JAN-2001 (first entry)

DE cDNA sequence of the wild type human FEZ1 gene.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX Homo sapiens.

OS WO200050565-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UUYE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Iehli H;

XX WPI; 2000-558396/51.

XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX Example 2; Fig 5B; 255pp; English.

XX The present sequence represents the cDNA sequence of the human FEZ1 gene.
CC FEZ1 is a tumour suppressor gene, located at chromosome location 8p22.
CC Decreased or no expression of FEZ1 is detected in a variety of cancer
CC cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.

XX Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;

Query Match 100.0%; Score 1791; DB 21; Length 5492;

Best Local Similarity 100.0%; Pred. No. 7.8e-288;

Matches 1791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGAGCGTCAGTAGCTCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTCCCGGCT 60
Db 112 ATGGGCGAGCGTCAGTAGCTCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTCCCGGCT 171
Qy 61 TCGCAGTACAAAGCTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCCACGGG 120
Db 172 TCGCAGTACAAAGCTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCCACGGG 231
Qy 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGC 180
Db 232 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGC 291
Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGAAAGCTCCGGGCTCCCATCACCA 240
Db 292 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGAAAGCTCCGGGCTCCCATCACCA 351
Qy 241 GATTACAGGCACTGTTCAGCGGGGATTAGGGGCGCAGGCTGGGGTGGACTTTGACCGG 300
Db 352 GATTACAGGCACTGTTCAGCGGGGATTAGGGGCGCAGGCTGGGGTGGACTTTGACCGG 411
Qy 301 TCCACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGT 360
Db 412 TCCACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGT 471
Qy 361 CGAGTAGGGGCCACACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTCTGACTCC 420
Db 472 CGAGTAGGGGCCACACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTCTGACTCC 531
Qy 421 TCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 480
Db 532 TCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 591
Qy 481 GAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGCGCGGAACTCCATGTCC 540
Db 592 GAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGCGCGGAACTCCATGTCC 651
Qy 541 AGCCTGCCACACACAGCAGCAGCAGCAGCTACAGCTGGAGCCCGCTGGTCCACACCCGTG 600
Db 652 AGCCTGCCACACACAGCAGCAGCAGCAGCTACAGCTGGAGCCCGCTGGTCCACACCCGTG 711
Qy 601 GGACCCCAAGCCGTTTGGGGGCTTCGCCCCACAAACATCACCCAGGGCATCGTCTCCAG 660
Db 712 GGACCCCAAGCCGTTTGGGGGCTTCGCCCCACAAACATCACCCAGGGCATCGTCTCCAG 771
Qy 661 GACAGCAACATGATGAGCCTGAAGCCTGTGCTTCTCCGAGGAGGTAGCAAGCTGGGC 720
Db 772 GACAGCAACATGATGAGCCTGAAGCCTGTGCTTCTCCGAGCAGGAGTAGCAAGCTGGGC 831
Qy 721 CACTCGAACAAGGAGCAGCAAGGGGCCCTCGTGTGTGCTCGCTCCCCCATCTCCACGGAGCAG 780
Db 832 CACTCGAACAAGGAGCAGCAAGGGGCCCTCGTGTGTGCTCGCTCCCCCATCTCCACGGAGCAG 891
Qy 781 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG 840
Db 892 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG 951
Qy 841 CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTGGGCTAGGAGAGGGCGGCGG 900
Db 952 CAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTGGGCTAGGAGAGGGCGGCGG 1011
Qy 901 CGCTGCGAGGACAGCTGGAGGGGCCCGAGGCCCAAGAGCGGCAACAGCTCAAGCAGAGCC 960
Db 1012 CGCTGCGAGGACAGCTGGAGGGGCCCGAGGCCCAAGAGCGGCAACAGCTCAAGCAGAGCC 1071
Qy 961 TCGCAGAAAGCCAGCGCGCAGCAGGTCTCTGACCTTGCAGGTACTGACAGTTCAGCAG 1020
Db 1072 TCGCAGAAAGCCAGCGCGCAGCAGGTCTCTGACCTTGCAGGTACTGACAGTTCAGCAG 1131
Qy 1021 GAGAAGCGGCGAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGAGCAGGAGCTTGTGAG 1080
Db 1132 GAGAAGCGGCGAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGAGCAGGAGCTTGTGAG 1191

QY 1081 ACCAAGCTCAGTCTTACGAGAGGAGAACACCAGCTTCGGCCCCCGCTCGAGAGACC 1140
 Db 1192 ACCAAGCTCAGTCTTACGAGAGGAGAGAACACCAGCTTCGGCCCCCGCTCGAGAGACC 1251
 QY 1141 CAGTGGAGGCTGTCGAGAGGTCAGGAGGATCTCCCTCTGAGAGCAGCAGCTGAAGGAG 1200
 Db 1252 CAGTGGAGGCTGTCGAGAGGTCAGGAGGATCTCCCTCTGAGAGCAGCAGCTGAAGGAG 1311
 QY 1201 TCCACAGCGAGGTGAACCCCAAGCTAGCAGAGATCTCTGGTCTCAAGCCACAGCTGAAG 1260
 Db 1312 TCCACAGCGAGGTGAACCCCAAGCTAGCAGAGATCTCTGGTCTCAAGCCACAGCTGAAG 1371
 QY 1261 GACACGGGGCCAGCTGGAGGCTGAGCTGAGGACCCAGGACCTGGAGGGCCCTG 1320
 Db 1372 GACACGGGGCCAGCTGGAGGCTGAGCTGAGGACCCAGGACCTGGAGGGCCCTG 1431
 QY 1321 CGACCAAGGGCTGAGCTGAGGCTGAGGATGAGCTGAGGACCAAGAAAGCAG 1380
 Db 1432 CGACCAAGGGCTGAGCTGAGGCTGAGGATGAGCTGAGGACCAAGAAAGCAG 1491
 QY 1381 GCGAGCTGCTCGGAGAGAGGTGAACCTGCTGGAGCAGAGCTGAGGAGCTGCGGCC 1440
 Db 1492 GCGAGCTGCTCGGAGAGAGGTGAACCTGCTGGAGCAGAGCTGAGGAGCTGCGGCC 1551
 QY 1441 CAGCGCGCCCTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1500
 Db 1552 CAGCGCGCCCTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1611
 QY 1501 CAGCGGAGCTGGAGCGCTGCGGCCGAGCTGCGGGAGGAGCGCAAGGCCATGACCAG 1560
 Db 1612 CAGCGGAGCTGGAGCGCTGCGGCCGAGCTGCGGGAGGAGCGCAAGGCCATGACCAG 1671
 QY 1561 ATGCTCTCGGCTTCCAGCATGAGCGCTCGTGTGGAAGAGGAGGAAGAGAGTGAAT 1620
 Db 1672 ATGCTCTCGGCTTCCAGCATGAGCGCTCGTGTGGAAGAGGAGGAAGAGAGTGAAT 1731
 QY 1621 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTACCAGCGAAACCGGCTG 1680
 Db 1732 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTACCAGCGAAACCGGCTG 1791
 QY 1681 GAGAGGGCCCTGACAGCTGCGAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1740
 Db 1792 GAGAGGGCCCTGACAGCTGCGAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1851
 QY 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791
 Db 1852 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1902

RESULT 3

AAAG4515
 ID AAAG4515 standard; cDNA; 1722 BP.

XX AC AAAG4515;

XX DT 02-JAN-2001 (first entry)

XX DE Nucleotide sequence of truncated FEZ1 transcript G3612.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
 KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;
 KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
 KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
 KW tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 XX CDS 1..1722
 XX /tag= a
 XX /product= "truncated FEZ1"

PN WO2000050565-A2.
 XX 31-AUG-2000.
 XX 25-FEB-2000; 2000WO-US04950.
 XX 25-FEB-1999; 99US-0121537.
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 XX Croce CM, Ishii H;
 XX WPI: 2000-558396/51.
 XX P-PSDB; AAB08722.
 DR New polynucleotide homologous with a portion of one strand of the human
 PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
 PT cancer -
 XX Disclosure; Fig 5H; 255pp; English.
 XX The present sequence encodes a truncated human FEZ1 polypeptide. The
 CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
 CC suppressor gene, located at chromosome location 8p22. Decreased
 CC or no expression of FEZ1 is detected in a variety of cancer cells.
 CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
 CC also interacts with tubulin, with microtubules, and with protein
 CC EPI-gamma. Post-translational phosphorylation and dephosphorylation
 CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
 CC expression are useful for inducing cells to proliferate. Compounds
 CC which modulate FEZ1 association with tubulin are useful for alleviating
 CC tubulin hyper- or hypo- polymerisation disorders, such as those
 CC associated with aberrant initiation of mitosis, modulation of the
 CC initiation and rate of cell proliferation and cell growth, modulation of
 CC cell shape, cell rigidity, cell motility, rate and stage of cellular
 CC DNA replication, intracellular distribution of organelles, metastatic
 CC potential of cell and cellular transformation from a non-cancerous to
 CC cancerous phenotype. Compounds which modulate FEZ1 binding and
 CC phosphorylation are also useful for alleviating a disorder, such as
 CC tumorigenesis, tumour survival, growth and metastasis.
 XX Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;
 XX
 Query Match 91.78; Score 1643; DB 21; Length 1722;
 Best Local Similarity 96.1%; Pred. No. 2.3e-263;
 Matches 1722; Conservative 0; Mismatches 0; Indels 69; Gaps 1;
 QY 1 ATGGGCGAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60
 Db 1 ATGGGCGAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60
 QY 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120
 Db 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120
 QY 121 CTGCTGAGGTTGGCTTCTCCAGAGCTCCGGTCACGCGAGTCCAGTCCAAATGGC 180
 Db 121 CTGCTGAGGTTGGCTTCTCCAGAGCTCCGGTCACGCGAGTCCAGTCCAAATGGC 180
 QY 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGCAAGAGCCCGGGGTCCCATCACCCA 240
 Db 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGCAAGAGCCCGGGGTCCCATCACCCA 240
 QY 241 GATTACAGGCACTGTCTCAGCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 300
 Db 241 GATTACAGGCACTGTCTCAGCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 300
 QY 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGGGCTCCGAGAGG 360
 Db 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGGGCTCCGAGAGG 360
 QY 361 CGAGTGAGGGCCACAGCGCTTCAAGCCTGTGTCTGCTCAGCGGTGAGGAGCCATCTGCACTCC 420

Db 361 GAGGTGAGGCGCACAGCCTTCAAGCCTGTGTGTCACCGGTGAGGAGCCATCTGCACTCC 420
Qy 421 TCCCGGAGAGTGCAGCAGCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 480
Db 421 TCCCGGAGAGTGCAGCAGCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 480
Qy 481 GAGCTGAAGCCTGGCCTGTGCTCTGGGGCGGTGTGAGACTCCCGCGGAACTCCATGTCC 540
Db 481 GAGCTGAAGCCTGGCCTGTGCTCTGGGGCGGTGTGAGACTCCCGCGGAACTCCATGTCC 540
Qy 541 AGCTGCCCCACACACAGCAGCAGCAGCTACACAGCTGGACCCGCTGGTCAACCCCGTG 600
Db 541 AGCTGCCCCACACACAGCAGCAGCAGCTACACAGCTGGACCCGCTGGTCAACCCCGTG 600
Qy 601 GGACCCACAAGCCGTTTTGGGGGCTCCGCCACACATCACCAGGGGATCGTCTCCAG 660
Db 601 GGACCCACAAGCCGTTTTGGGGGCTCCGCCACACATCACCAGGGGATCGTCTCCAG 660
Qy 661 GACAGCAATGATGAGCCTGAGGCTCTGTCTTCTCCGACGGAGGTAGCAAGCTGGGC 720
Db 661 GACAGCAATGATGAGCCTGAGGCTCTGTCTTCTCCGACGGAGGTAGCAAGCTGGGC 720
Qy 721 CACTCGACAGGAGCTGGAGCAGAACTGTGTCGCTCCCTCCATCTCCACGAGCAG 780
Db 721 CACTCGACAGGAGCTGGAGCAGAACTGTGTCGCTCCCTCCATCTCCACGAGCAG 780
Qy 781 TGCAGATCCAGGAGCTGGAGCAGAACTGTGTCGCTCCCTCCATCTCCACGAGCAG 840
Db 781 TGCAGATCCAGGAGCTGGAGCAGAACTGTGTCGCTCCCTCCATCTCCACGAGCAG 840
Qy 841 CAGCGACCTTTGAGGAGAGAGCTTGCCTCCAGCCTGGCTACGAGAGCGCGCGG 900
Db 841 CAGCGACCTTTGAGGAGAGAGCTTGCCTCCAGCCTGGCTACGAGAGCGCGCGG 900
Qy 901 CGCTCGAGGAGAGCTGGAGGCGCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCC 960
Db 901 CGCTCGAGGAGAGCTGGAGGCGCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCC 960
Qy 961 TCGCAGAGAGCAGCGCGCGCAGAGCTCTGCACCTGCAGGTACTGCAGCTTCAGCAG 1020
Db 961 TCGCAGAGAGCAGCGCGCGCAGAGCTCTGCACCTGCAGGTACTGCAGCTTCAGCAG 1020
Qy 1021 GAGAAAGCGGAGCTCCGCAAGAGCTCGAGAGCTCATGAAGAGCAGGACCTGCTGAG 1080
Db 1021 GAGAAAGCGGAGCTCCGCAAGAGCTCGAGAGCTCATGAAGAGCAGGACCTGCTGAG 1080
Qy 1081 ACCAGCTCAGGTCTTACAGAGGAGAGAACAGCAGCTTCGGCCCGCGCTGGAGAGACC 1140
Db 1081 ACCAGCTCAGGTCTTACAGAGGAGAGAACAGCAGCTTCGGCCCGCGCTGGAGAGACC 1140
Qy 1141 CAGTGGAGGTGTGCCAGAGTCAAGGAGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200
Db 1141 CAGTGGAGGTGTGCCAGAGTCAAGGAGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200
Qy 1201 TCCAGACGAGGTGAACCCAGAGCTAGCAGATCTCTGGGTCTCAAGGCAAGCTGAAG 1260
Db 1201 TCCAGACGAGGTGAACCCAGAGCTAGCAGATCTCTGGGTCTCAAGGCAAGCTGAAG 1260
Qy 1261 GACACCGCGGCAAGCTGGAGGCTTGGAGTGGAGACCCAGGACCTGGAGGGCGCCCTG 1320
Db 1261 GACACCGCGGCAAGCTGGAGGCTTGGAGTGGAGACCCAGGACCTGGAGGGCGCCCTG 1320
Qy 1321 CGCACAGGCGCTGAGGCTGTGAGATGAGTGCAGCGCAAGAAAGCAG 1380
Db 1321 CGCACAGGCGCTGAGGCTGTGAGATGAGTGCAGCGCAAGAAAGCAG 1380
Qy 1381 GCGAGCTGCTCGGAGAGAGGTGAACCTGTGGAGAGGAGCTGCAGAGCTGCGGCCC 1440
Db 1381 GCGAGCTGCTCGGAGAGAGGTGAACCTGTGGAGAGGAGCTGCAGAGCTGCGGCCC 1440
Qy 1441 CAGCGCGCTGCGCGGAGCATGAGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCGCTG 1500
Db 1441 CAGCGCGCTGCGCGGAGCATGAGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCGCTG 1500

Qy 1501 CAGCGGAGCTGGAGCGGCTCGGCGCGAGCTGCGGGAGGAGCGCAAGGCCATGACCAG 1560
Db 1501 CAGCGGAGCTG----- 1512
Qy 1561 ATGCTCTCGGCTTCCAGCATGAGCGGCTCTGTGGAAAGGAGGAGAAAGGAGGTGATT 1620
Db 1513 -----GAGCGGCTCTGTGGAAAGGAGGAGAAAGGAGGTGATT 1551
Qy 1621 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACACGCGAAACAGCGCCTG 1680
Db 1552 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACACGCGAAACAGCGCCTG 1611
Qy 1681 GAGAAGGCCCTGCAGCAGCTGGCGACAGCGCGGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1740
Db 1612 GAGAAGGCCCTGCAGCAGCTGGCGACAGCGCGGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1671
Qy 1741 CTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1791
Db 1672 CTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1722

RESULT 4
AAA64514

ID AAA64514 standard; cDNA; 1692 BP.

XX AAA64514;

AC AC

DT 02-JAN-2001 (first entry)

XX Nucleotide sequence of truncated FEZ1 transcript G3611.

Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
tumour proliferation; tubulin; microtubule; protein EF1-gamma;
tubulin polymerisation disorder; mitosis initiation; cell proliferation;
cell growth; cell shape; cell rigidity; cell motility; DNA replication;
tumorigenesis; tumour survival; metastasis; ss.

XX Homo sapiens.

Key Location/Qualifiers
CDS 1..1692

FT /*tag= a

FT /product= "truncated FEZ1"

XX WO200005065-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Ishii H;

XX WPI; 2000-558396/51.

XX P-PSDB; AAB08721.

XX New polynucleotide homologous with a portion of one strand of the human
FEZ1 gene, useful for alleviating abnormal cell proliferation such as
cancer -

XX Disclosure; Fig 5G; 255pp; English.

XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EF1-gamma. Post-translational phosphorylation and dephosphorylation

modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are useful for inducing cells to proliferate. Compounds which modulate FEZ1 association with tubulin are useful for alleviating tubulin hyper- or hypo- polymerisation disorders, such as those associated with aberrant initiation of mitosis, modulation of the initiation and rate of cell proliferation and cell growth, modulation of cell shape, cell rigidity, cell motility, rate and stage of cellular DNA replication, intracellular distribution of organelles, metastatic potential of cell and cellular transformation from a non-cancerous to cancerous phenotype. Compounds which modulate FEZ1 binding and phosphorylation are also useful for alleviating a disorder, such as tumorigenesis, tumour survival, growth and metastasis.

SQ Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;

Query Match 88.3%; Score 1581.4; DB 21; Length 1692;
 Best Local Similarity 94.4%; Pred. No. 3.7e-253;
 Matches 1691; Conservative 0; Mismatches 1; Indels 99; Gaps 1;

QY 1 ATGGCAGGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60
 DB 1 ATGGCAGGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60

QY 61 TCGCAGTACAAGTCGCGAAGTCCCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 120
 DB 61 TCGCAGTACAAGTCGCGAAGTCCCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 120

QY 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGTCCCAAAATGGGC 180
 DB 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGTCCCAAAATGGGC 180

QY 181 AAGCGAAGACTTCTTCTACATCAAGTCAAGCAAGTCCAGTCCCAAAATGGGC 240
 DB 181 AAGCGAAGACTTCTTCTACATCAAGTCAAGCAAGTCCAGTCCCAAAATGGGC 240

QY 241 GATTACAGGCACTCTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGATTTGACCCG 300
 DB 241 GATTACAGGCACTCTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGATTTGACCCG 300

QY 301 TCCACACCCCAAGCTCATGCTTCTCCAAATAGTAAATGGGTCCGAGAGGGT 360
 DB 301 TCCACACCCCAAGCTCATGCTTCTCCAAATAGTAAATGGGTCCGAGAGGGT 360

QY 361 GCAGTGAGGCCACAGCTTCAAGCTTGCTGCCACCGTCAAGGAGCCATCTGCACCTCC 420
 DB 361 GCAGTGAGGCCACAGCTTCAAGCTTGCTGCCACCGTCAAGGAGCCATCTGCACCTCC 420

QY 421 TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGCAAGCCCAAGGAGCAG 480
 DB 421 TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGCAAGCCCAAGGAGCAG 480

QY 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCTGTGCTCAGACTCCGGCCGGAATCCATGTCC 540
 DB 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCTGTGCTCAGACTCCGGCCGGAATCCATGTCC 540

QY 541 AGCTGCCACACACAGCAGCAGCAGCTACAGCTGAGCTGGACCCCTGGTCAACCCCTG 600
 DB 541 AGCTGCCACACACAGCAGCAGCAGCTACAGCTGAGCTGGACCCCTGGTCAACCCCTG 600

QY 601 GGACCCACAAAGCCGTTTGGGGCTCCCGCCCAACATCACCCAGGGATCGTCTCCAG 660
 DB 601 GGACCCACAAAGCCGTTTGGGGCTCCCGCCCAACATCACCCAGGGATCGTCTCCAG 660

QY 661 GACAGCAATGATGAGCTGAAGCTGTCTCTTCTCCGAGAGGTAGCAAGCTGGGC 720
 DB 661 GACAGCAATGATGAGCTGAAGCTGTCTCTTCTCCGAGAGGTAGCAAGCTGGGC 720

QY 721 CACTCGAACAGCAGCAGAGGCCCTCGTGTGCTGGCTCCCTCCCTCCACGACGAG 780
 DB 721 CACTCGAACAGCAGCAGAGGCCCTCGTGTGCTGGCTCCCTCCCTCCACGACGAG 780

QY 781 TGCAGCATCCAGGAGCTGAGCAGAGCTGTTGGAGAGGGGGCGCCCTCCAGAAAGCTG 840
 DB 781 TGCAGCATCCAGGAGCTGAGCAGAGCTGTTGGAGAGGGGGCGCCCTCCAGAAAGCTG 840

DB 781 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTTGGAGAGGGGGCGCCCTCCAGAAAGCTG 840
 QY 841 CAGCGCAGCTTTGAGGAGAGAGAGCTTTCCTCCAGCTGCGCTTACGAGGAGCGCGCGG 900
 DB 841 CAGCGCAGCTTTGAGGAGAGAGAGCTTTCCTCCAGCTGCGCTTACGAGGAGCGCGCGG 900
 QY 901 CGCTGACGAGCAGCAGCTGAGAGGGCGCGAGCCCAAGCGCGGCAACAAAGCTCAAGCAGGCC 960
 DB 901 CGCTGACGAGCAGCAGCTGAGAGGGCGCGAGCCCAAGCGCGGCAACAAAGCTCAAGCAGGCC 960
 QY 961 TCGCAGAAAGCAGCAGCGCGCAGCAGCTCTGACCTGACCTGAGGTACTGACGCTTTCAGCAG 1020
 DB 961 TCGCAGAAAGCAGCAGCGCGCAGCAGCTCTGACCTGACCTGAGGTACTGACGCTTTCAGCAG 1020

QY 1021 GAGAAAGCAGCTCCCGCAGGAGCTTCAGAGCTCATGAAGGAGCAGGACCTCTCTGGAG 1080
 DB 1021 GAGAAAGCAGCTCCCGCAGGAGCTTCAGAGCTCATGAAGGAGCAGGACCTCTCTGGAG 1080

QY 1081 ACCAAGCTTCAGGTCCTACGAGAGGGAGAAAGCAGCTTTCGGCCCGCGCTGGAGGAGACC 1140
 DB 1081 ACCAAGCTTCAGGTCCTACGAGAGGGAGAAAGCAGCTTTCGGCCCGCGCTGGAGGAGACC 1140

QY 1141 CAGTGGAGGTGTGCCAAGAGTCAAGGAGAGTCTTCCCTCTGAAAGCAGCAGCTGAAGGAG 1200
 DB 1141 CAGTGGAGGTGTGCCAAGAGTCAAGGAGAGTCTTCCCTCTGAAAGCAGCAGCTGAAGGAG 1200

QY 1201 TCCAGACGAGGTGAACCGCAAGGCTAGCGAGATCTTGGTCTCAAGGCAAGCTGAAG 1260
 DB 1201 TCCAGACGAGGTGAACCGCAAGGCTAGCGAGATCTTGGTCTCAAGGCAAGCTGAAG 1260

QY 1261 GACACGCGGGCAGCTGGAGGGCTCTGAGCTGAGGACCCAGGACCTTGGAGGGCGCCCTG 1320
 DB 1261 GACACGCGGGCAGCTGGAGGGCTCTGAGCTGAGGACCCAGGACCTTGGAGGGCGCCCTG 1320

QY 1321 CGCACCAAGGCGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGACGCGCAAGAAACGAG 1380
 DB 1321 CGCACCAAGGCGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGACGCGCAAGAAACGAG 1380

QY 1381 GCGAGCTGTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGCGAGGAGCTGGGGCC 1440
 DB 1381 GCGAGCTGTGCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGCGAGGAGCTGGGGCC 1440

QY 1441 CAGGCGCGCTGGCGCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCCCTG 1500
 DB 1441 CAGGCGCGCTGGCGCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCCCTG 1500

QY 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGAGCTGCGGAGGAGCGGCAAGGCAATGACGAG 1560
 DB 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGAGCTGCGGAGGAGCGGCAAGGCAATGACGAG 1560

QY 1561 ATGTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAGGAGGAGGAGGAGGAGGAGT 1620
 DB 1561 ATGTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAGGAGGAGGAGGAGGAGGAGT 1620

QY 1621 CAGTACCAAGAAACAGCTGACGAGCTGAGTGTGCGCATGTACCAAGCGGAAACAGCGCCTG 1680
 DB 1621 CAGTACCAAGAAACAGCTGACGAGCTGAGTGTGCGCATGTACCAAGCGGAAACAGCGCCTG 1680

QY 1681 GAGAGGCGCTTGCAGCAGCTGGCAGCTGGGAGCAGCGCGGGAGCGCCCTTCGAGGTTCAC 1740
 DB 1681 GAGAGGCGCTTGCAGCAGCTGGCAGCTGGGAGCAGCGCGGGAGCGCCCTTCGAGGTTCAC 1740

QY 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791
 DB 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791

QY 1841 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1891
 DB 1841 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1891

RESULT 5
 AAA64512
 ID AAA64512 standard; cDNA; 1614 BP.
 XX
 AC
 AC
 XX
 XX

ID	Accession	Title
02-JAN-2001	(first entry)	
Nucleotide sequence of truncated FEZ1 transcript T8D145M4.		
Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth; tumour proliferation; tubulin; microtubule; protein EFl-gamma; tubulin polymerisation disorder; mitosis initiation; cell proliferation; cell growth; cell shape; cell rigidity; cell motility; DNA replication; tumorigenesis; tumour survival; metastasis; ss.		
Homo sapiens.		
XX	Key	Location/Qualifiers
PH	CDS	1..633
FT	FT	/tag= a
FT	FT	/product= "truncated FEZ1"
W0200050565-A2.		
31-AUG-2000.		
25-FEB-2000; 2000MO-US04950.		
25-FEB-1999; 99US-0121537.		
(UYJE-) UNIV JEFFERSON THOMAS.		
Croce CM, Ishii H;		
WPI; 2000-558396/51.		
P-PSDB; AAB08719.		
New polynucleotide homologous with a portion of one strand of the human FEZ1 gene, useful for alleviating abnormal cell proliferation such as cancer -		
Disclosure; Fig 5E; 255pp; English.		
The present sequence encodes a truncated human FEZ1 polypeptide. The encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour suppressor gene, located at chromosome location 8p22. Decreased or no expression of FEZ1 is detected in a variety of cancer cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts with tubulin, with microtubules, and with protein EFl-gamma. Post-translational phosphorylation and dephosphorylation modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are useful for inducing cells to proliferate. Compounds which modulate FEZ1 association with tubulin are useful for alleviating tubulin hyper- or hypo- polymerisation disorders, such as those associated with aberrant initiation of mitosis, modulation of the initiation and rate of cell proliferation and cell growth, modulation of cell shape, cell rigidity, cell motility, rate and stage of cellular DNA replication, intracellular distribution of organelles, metastatic potential of cell and cellular transformation from a non-cancerous to cancerous phenotype. Compounds which modulate FEZ1 binding and phosphorylation are also useful for alleviating a disorder, such as tumorigenesis, tumour survival, growth and metastasis.		
Sequence 1614 BP; 367 A; 500 C; 522 G; 225 T; 0 other;		
Query Match	79.7%; Score 1427; DB 21; Length 1614;	
Best Local Similarity	90.1%; Pred. No. 1.3e-227;	
Matches 1614; Conservative	0; Mismatches 0; Indels 177; Gaps 1;	
QY	1	ATGGGAGCGCTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGCT 60
DB	1	ATGGGAGCGCTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGCT 60
QY	61	TGCGAGTACAAGCTGGCGAAGTCCTCCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
DB	61	TGCGAGTACAAGCTGGCGAAGTCCTCCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
QY	121	CTGCTGAGGTTTGCTTCTCCAGGACTCCGGTCAAGGCTCAGGCTCCAAATGGGC 180

Db	1486	CGGACCTCCATGCTCCAGCTGCCACACACAGACACAGACGACGCTACGAGCTGGACCC	1545
Qy	585	GCTGGTACACCGTGGACCCACAAAGCCGTTTGGGGGCTCCGCCACAAATCATCCCA	644
Db	1546	GCTGGTACACCGTGGACCCACAAAGCCGTTTGGGGGCTCCGCCACAAATCATCCCA	1605
Qy	645	GGGATCGTCTCCAGGACACACATGATGAGCTGAAGGCTGTCTCTCCGACGG	704
Db	1606	GGGATCGTCTCCAGGACACACATGATGAGCTGAAGGCTGTCTCTCCGACGG	1665
Qy	705	AGGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGCCCTCGTGTCCGCTCCCC	764
Db	1666	AGGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGCCCTCGTGTCCGCTCCCC	1725
Qy	765	CATCTCCAGCAGAGTCCAGGATCCAGGAGCTGGAGCAGAGCTGTTGGAGAGGAGGG	824
Db	1726	CATCTCCAGCAGAGTCCAGGATCCAGGAGCTGGAGCAGAGCTGTTGGAGAGGAGGG	1785
Qy	825	CGCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCCTGGCCTA	884
Db	1786	CGCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCCTGGCCTA	1845
Qy	885	CGAGAGCGCGCGCGCTCGAGGACGAGCTGGAGGCGCGGAGCCCAAGCGGCA	944
Db	1846	CGAGAGCGCGCGCGCTCGAGGACGAGCTGGAGGCGCGGAGCCCAAGCGGCA	1905
Qy	945	CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCAGCAGGTCCTGCACTGCAGGT	1004
Db	1906	CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCAGCAGGTCCTGCACTGCAGGT	1965
Qy	1005	ACTCAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCCGAGGCTCATCAGGA	1064
Db	1966	ACTCAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCCGAGGCTCATCAGGA	2025
Qy	1065	CGAGGACCTGTGGAGACCAAGCTCAGGCTTACGAGGAGGAGAGCCAGCTTCGGGCC	1124
Db	2026	CGAGGACCTGTGGAGACCAAGCTCAGGCTTACGAGGAGGAGAGCCAGCTTCGGGCC	2085
Qy	1125	CGCGTGGAGAGACCCAGTGGAGGTGGCC	1156
Db	2086	CGCGTGGAGAGACCCAGTGGAGGTGGCC	2117

RESULT 8

AA528699
ID AA528699 standard; DNA; 2411 BP.

AC AA528699;

XX 07-NOV-2001 (first entry)

DE Genomic sequence #539 encoding for novel human respiratory antigen.

XX Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active; ds.

XX Homo sapiens.

OS WO200155448-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01333.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0198874.

PR 17-MAR-2000; 2000US-0190076.

PR	18-APR-2000;	2000US-0198123.
PR	19-MAY-2000;	2000US-0205515.
PR	07-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226688.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	06-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233083.
PR	14-SEP-2000;	2000US-0233084.
PR	21-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235836.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.


```

Db      2086 CGCGCTGGAGGAGACCCAGTGGGAGGTGAGGC 2117
RESULT 9
AAA64507
ID      AAA64507 standard; DNA; 9048 BP.
XX
AC      AAA64507;
XX
DT      02-JAN-2001 (first entry)
XX
DE      Nucleotide sequence comprising the human FEZ1 gene.
XX
KW      Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW      tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW      tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW      cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW      tumorigenesis; tumour survival; metastasis; ss.
XX
OS      Homo sapiens.
XX
PN      WO2000050565-A2.
XX
PD      31-AUG-2000.
XX
PF      25-FEB-2000; 2000WO-US04950.
XX
PR      25-FEB-1999; 99US-0121537.
XX
PA      (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI      Croce CM, Ishii H;
XX
WPI, 2000-558396/51.
XX
New polynucleotide homologous with a portion of one strand of the human
FEZ1 gene, useful for alleviating abnormal cell proliferation such as
cancer -
XX
Claim 2; Fig 5A; 255pp; English.
XX
The present sequence comprises the human FEZ1 gene. FEZ1 is a tumour
suppressor gene, located at chromosome location 8p22. Decreased or no
expression of FEZ1 is detected in a variety of cancer cells. Expression
of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts
with tubulin, with microtubules, and with protein EPI-gamma.
XX
Post-translational phosphorylation and dephosphorylation modulates the
effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are
useful for inducing cells to proliferate. Compounds which modulate FEZ1
association with tubulin are useful for alleviating tubulin hyper- or
hypo- polymerisation disorders, such as those associated with aberrant
initiation of mitosis, modulation of the initiation and rate of cell
proliferation and cell growth, modulation of cell shape, cell rigidity,
cell motility, rate and stage of cellular DNA replication, intracellular
distribution of organelles, metastatic potential of cell and cellular
transformation from a non-cancerous to cancerous phenotype. Compounds
which modulate FEZ1 binding and phosphorylation are also useful for
alleviating a disorder, such as tumorigenesis, tumour survival, growth
and metastasis.
XX
SQ      Sequence 9048 BP; 2011 A; 2605 C; 2490 G; 1939 T; 3 other;
Query Match      45.1%; Score 807.2; DB 21; Length 9048;
Best Local Similarity 99.6%; Pred. No. 5.8e-125;
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
345 GGGCTCCGAGAGGGTGCAGTGAGGCCACAGCTTCAAGCTGTGCTGCCACCGTCCAGG 404
1646 GGGCTCCGAGAGGGTGCAGTGAGGCCACAGCTTCAAGCTGTGCTGCCACCGTCCAGG 1705
XX
405 AGCCATCTGCACTCTCTCCCGGAGTGCCAGCCACAGCTGCACCCGCCCTCCAGA 464
1706 AGCCATCTGCACTCTCTCCCGGAGTGCCAGCCACAGCTGCACCCGCCCTCCAGA 1765
XX
465 CAAGCCCAAGGAGCAGCAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 524
1766 CAAGCCCAAGGAGCAGCAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 1825
XX
525 CCGGAATCTCCATGTCCAGCTGCCCCACACAGCAGCAGCAGCAGCTACCACTGGACCC 584
1826 CCGGAATCTCCATGTCCAGCTGCCCCACACAGCAGCAGCAGCAGCTACCACTGGACCC 1885
XX
585 GCTGGTCACACCCGTGGGACCCACAAGCGTTTGGGGGCTCCGCCCAACATCACCAC 644
1886 GCTGGTCACACCCGTGGGACCCACAAGCGTTTGGGGGCTCCGCCCAACATCACCAC 1945
XX
645 GGGCATCGTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTCTTCTCCGACGG 704
1946 GGGCATCGTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTCTTCTCCGACGG 2005
XX
705 AGGTAGCAAGCTGGGCGCACTCGAAACAAGGAGCAGCAAGGGCCCTCGTGTGCTCCGC 764
2006 AGGTAGCAAGCTGGGCGCACTCGAAACAAGGAGCAGCAAGGGCCCTCGTGTGCTCCGC 2065
XX
765 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAGCAGCAAGCTGTGAGAGGGAGGG 824
2066 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAAACAAGGAGCTGTGAGAGGGAGGG 2125
XX
825 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGAGCTTGCCTCCAGCTGGCCTA 884
2126 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGAGCTTGCCTCCAGCTGGCCTA 2185
XX
885 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGAGGAGCGCGGAGCCCAAGGCGGCAA 944
2186 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGAGGAGCGCGGAGCCCAAGGCGGCAA 2245
XX
945 CAAGCTCAAGCAGCGCTCGCAGAGAGCAGCCGCGGCGCAGCAGCTCTGCACTGCAGGT 1004
2246 CAAGCTCAAGCAGCGCTCGCAGAGAGCAGCCGCGGCGCAGCAGCTCTGCACTGCAGGT 2305
XX
1005 ACTCAGCTTCCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGA 1064
2306 ACTCAGCTTCCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGA 2365
XX
1065 GCAGGACCTGTGGAGACCAAGCTCAGTCTTACGAGAGGAGAGAACAGCTTCGGCCC 1124
2366 GCAGGACCTGTGGAGACCAAGCTCAGTCTTACGAGAGGAGAGAACAGCTTCGGCCC 2425
XX
1125 CGCGCTGGAGAGACCCAGTGGGAGGTGTC 1156
2426 CGCGCTGGAGAGACCCAGTGGGAGGTGAGGC 2457
XX
RESULT 10
AAA64511
ID      AAA64511 standard; cDNA; 633 BP.
XX
AC      AAA64511;
XX
DT      02-JAN-2001 (first entry)
XX
DE      Nucleotide sequence of truncated FEZ1 transcript E264162.
XX
KW      Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW      tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW      tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW      cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW      tumorigenesis; tumour survival; metastasis; ss.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..633
/*tag= a
/product= "truncated FEZ1"
XX
```


29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
01-NOV-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249219.
17-NOV-2000; 2000US-0249220.
17-NOV-2000; 2000US-0249221.
17-NOV-2000; 2000US-0249222.
17-NOV-2000; 2000US-0249223.
17-NOV-2000; 2000US-0249224.
17-NOV-2000; 2000US-0249225.
17-NOV-2000; 2000US-0249226.
17-NOV-2000; 2000US-0249227.
17-NOV-2000; 2000US-0249228.
17-NOV-2000; 2000US-0249229.
17-NOV-2000; 2000US-0249230.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0251989.
06-DEC-2000; 2000US-0251990.
06-DEC-2000; 2000US-0251991.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251857.
08-DEC-2000; 2000US-0251858.
08-DEC-2000; 2000US-0251859.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-476224/51.
P-PSDB; AAU17943.

XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory
PT cancers and also for testing and detection e.g. diagnosis -
XX Claim 4; SED ID No 269; 546pp; English.
XX The present invention relates to the isolation of novel human
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful for preventing, treating and/or prognosing
CC disorders related to the respiratory system including throat
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS27869-AAS28159 encode for novel human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 560 BP; 120 A; 136 C; 161 G; 83 T; 0 other;
Query Match 31.3%; Score 560; DB 22; Length 560;
Best Local Similarity 100.0%; Pred. NO. 4.3e-84;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 311 CCAAGCTCATGCCCTTCTCCCAATCAGCTAGTAATGGGCTCCGAGAGGGTGCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCCAATCAGCTAGTAATGGGCTCCGAGAGGGTGCAGTGAGGC 60
Qy 371 CCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTGCACCTCTCCCGGAGA 430
Db . 61 CCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTGCACCTCTCCCGGAGA 120
Qy 431 GTGCCAGCCACCAGCTGCACCCCGCCCTCCAGACAAAGCCCAAGGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACCAGCTGCACCCCGCCCTCCAGACAAAGCCCAAGGAGCAGAGCTGAAGC 180
Qy 491 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGCGCGGAATCTCATGTCCAGCTGCCCA 550
Db 181 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGCGCGGAATCTCATGTCCAGCTGCCCA 240
Qy 551 CACACAGCACCAGCAGCAGCTACCAAGCTGGACCCGCTGGTCCACACCCGTGGGACCCACAA 610
Db 241 CACACAGCACCAGCAGCAGCTACCAAGCTGGACCCGCTGGTCCACACCCGTGGGACCCACAA 300
Qy 611 GCCGTTTTGGGGGCTCCGCCCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 670
Db 301 GCCGTTTTGGGGGCTCCGCCCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 360
Qy 671 TGATGAGCCTGAAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGCCACTCGAACA 730
Db 361 TGATGAGCCTGAAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGCCACTCGAACA 420
Qy 731 AGGCAGACAAGGGCCCTCGTGTGCTCGCTCCCGATCTCCACGAGCAGGTGCAGCATCC 790
Db 421 AGGCAGACAAGGGCCCTCGTGTGCTCGCTCCCGATCTCCACGAGCAGGTGCAGCATCC 480
Qy 791 AGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGGCAGCT 850
Db 481 AGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGGCAGCT 540
Qy 851 TTGAGGAGAGGAGCTTGCC 870
Db 541 TTGAGGAGAGGAGCTTGCC 560
RESULT 13
ABQ54361
ID ABQ54361 standard; cDNA; 1623 BP.
XX

XX DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:60.
XX DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antinefactive; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX OS Homo sapiens.
XX OS WO200055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
XX DR P-PSDB; AAB56422.
XX PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX Claim 1; Page 657; 2338pp; English.
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antinefactive, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX SQ Sequence 1617 BP; 355 A; 457 C; 532 G; 268 T; 5 other;
Query Match 14.4%; Score 258.2; DB 21; Length 1617;
Best Local Similarity 60.2%; Pred No. 4.1e-34;
Matches 523; Conservative 0; Mismatches 303; Indels 43; Gaps 4;
QY 962 CCAGAGAGCCAGCGCGCAGCAGGCTCTGACCTGACGTACTGAGCTTCAGCAGG 1021
Db 33 CACAGCGGCGACAGCGGCGCCCAAGAGTGTGACGTGACGTGTCCAGCTGCA-CAGG 91
QY 1022 AGAAGCGGCGAGCTCGGCGAGGCTGAGAGCTCATGAGGAGCAGGACCTGCTGGAGA 1081
Db 92 AGAAGCGGCAATTGAGGAGCAGCTTCGACAGCTGCTGAGGAGCGGCAACAGCTGGAGC 151
QY 1082 CCAAGCTCAGGTCTTACAGAGGAGAGACAGCAGCTTCGGCCCGCTGGAGGAGCCC 1141
Db 152 GCGCGTCCGCCACCTTGGAGCGGAGACAGCGGAGC-TCGGCGCGAGGCTTGGAGGACCA 210
QY 1142 AGTGGAGGTGTGCCAGAGTTCAGCGGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGGT 1201
Db 211 AGTGGAGGTGTGCCAGAGTTCAGCGGAGATCTCCCTGCTGAAGCAGCAGCTGAAGAGT 270
QY 1202 CCCAGACGGAGGTGAACCGCAGAGGTACGAGATCCTGGGTCTCAAGGCACAGCTGAAGG 1261

Db 271 CTCAGGCGAGCTGGTGCAGAAAGGCGCAGCTGGTGGCTCTCGCGGTGCGGTGCGGG 330
QY 1362 ACAGCGCGGCAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCGCTTCG 1321
Db 331 AGGCGCGTGTACCTGCGGTGAGTGGGCGCTGCGGCGGTCTACAGAGGCGCGCC 390
QY 1322 GCACAAAGGCGCTGGAGCTGGAGTCTGTGAGATGAGCTCAGCGCAAGAAACGAGG 1381
Db 391 GAGCTCGGAGCTGGAGCTGGAAGCTGTTCCAGGAGCTCAGCGCACCGCCAGGAG 450
QY 1382 CGGAGCTGCTCGGCGGAGAGTGAACCTGCTGGAGCAGGAGCTCAGAGCTGCGGG--- 1438
Db 451 CTGAGCAGCTCGGCGGAGAAAGCTGGGAGTTGATGCTGGCGCGCGGAGCTCCGGGAGC 510
QY 1439 -----CCAGGCGCGCTGCGCGCGGACATGG 1465
Db 511 CCCCTGTGCCACCTGCCACCGCTGACCCATTCCTCTGCGAGAGATGATGAGGCCAAG 570
QY 1466 GCGCGCCACCTTCCCGGAGGAGCTGCTGCTGCGAGCGGAGCTGGAGCGGTGCGGG 1525
Db 571 TGCAGCGGCGCAGCAGCGCGGTGCGGCGCAGCTTGGGCGCAGCTGAGCGATTTGCGGG 630
QY 1526 CCGAGCTCGGCGGAGCGGCAAGCCATGACCATGCTCTCGGGCTTCAGCATGAGC 1585
Db 631 TGGAGCTGCGAGCGGAGCGCGCGGTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGC 690
QY 1586 GCGCTGTGGAAGGAGGAGAGAGAGAGGAGTTCAGTACAGAAACAGCTGCAGCAGA 1645
Db 691 GCGTGGCTGCGAGCAGAGAGAGAGAGGAGTTCGCTACAGAGAGCTGCAGCACA 750
QY 1646 GCTAGTGGCCATGTACCGCGGAAACCGGCGAGTTCGAGAGAGCGCTGCGAGCTGGCAC 1705
Db 751 ACTACATCCAGATGTACCGCGGCAACCGGCGAGTTCAGCAGGAGCTGCGAGCTCAGCC 810
QY 1706 GTGCGGAGCAGCGCGGCGGAGCGCTTGGAGGTGACCTG-----GAGGGGCTGACATCCC 1760
Db 811 TGGAGCTGGAGCGCGGCGGAGCTTCGCTGACCTGGGCGCTGCGCGAGCGCGCTTCGATCTG 870
QY 1761 CTACGAGGACATCATGCGCACTGAGATCT 1789
Db 871 CCTGAGGAGATCACTGCTACTGAGATCT 899
RESULT 15
AAA64510
ID AAA64510 standard; cDNA; 404 BP.
XX AC AAA64510;
XX DT 02-JAN-2001 (first entry)
XX DE Nucleotide sequence of truncated FEZ1 transcript E16T8.
XX KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein B1-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..231
XX FT /*tag= a
XX FT /product= "truncated FEZ1"
XX PN WO200050565-A2.
XX PD 31-AUG-2000.
XX PF 25-FEB-2000; 2000WO-US04950.
XX PR 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.
XX PA Croce CM, Iehii H;
XX PI WPI; 2000-558396/51.
XX DR P-PSDB; AAB08717.
XX DR
XX
XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
XX Disclosure; Fig 5C; 255pp; English.
XX
XX The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EFl-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;

Query Match 13.9%; Score 249; DB 21; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.3e-32;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1543 CGGCAAGGCCCATGACCATGTCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGAAGGAG 1602
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
156 CGGCAAGGCCCATGACCATGTCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGAAGGAG 215
QY 1603 GAGAGGAGAGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGAGCTAGTGGCCATGTAC 1662
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
216 GAGAGGAGAGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGAGCTAGTGGCCATGTAC 275
QY 1663 CAGCGGAACCCAGCGCTCGAGAGGCCCTCGAGCAGCTGGGACAGCGCCCGG 1722
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
276 CAGCGGAACCCAGCGCTCGAGAGGCCCTCGAGCAGCTGGGACAGCGCCCGG 335
QY 1723 GAGCCCTTTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACT 1782
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
336 GAGCCCTTTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACT 395
QY 1783 GAGATCTGA 1791
DB ||||||||
396 GAGATCTGA 404

Search completed: June 14, 2003, 23:43:13
Job time : 419.345 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	146	8.2	1926	4	US-09-249-585A-2	Sequence 2, Appli
2	146	8.2	2580	3	US-09-050-861-2	Sequence 2, Appli
3	146	8.2	2580	4	US-09-359-083-2	Sequence 2, Appli
C 4	146	8.2	5452	2	US-09-130-114-1	Sequence 1, Appli
5	146	8.2	9600	4	US-08-910-647-1	Sequence 1, Appli
6	146	8.2	9600	4	US-09-620-925-1	Sequence 1, Appli
7	146	8.2	10596	1	US-07-884-811-15	Sequence 15, Appl
8	146	8.2	10596	1	US-07-885-971-15	Sequence 15, Appl
9	146	8.2	10596	1	US-08-087-783A-15	Sequence 15, Appl
10	146	8.2	10596	1	US-08-194-088B-15	Sequence 15, Appl
11	146	8.2	10596	2	US-08-194-087-15	Sequence 15, Appl
12	146	8.2	10596	5	PCR-US93-04648-15	Sequence 15, Appl
13	138.4	7.7	3489	2	US-08-728-323A-1	Sequence 1, Appli
14	138.4	7.7	3489	4	US-09-298-568-1	Sequence 1, Appli
C 15	138.4	7.7	32207	2	US-08-770-379-20	Sequence 20, Appl
C 16	138.4	7.7	32207	4	US-08-757-669A-20	Sequence 20, Appl
C 17	138.4	7.7	32207	4	US-09-230-371A-20	Sequence 20, Appl
18	125.2	7.0	9551	1	US-08-056-200-93	Sequence 93, Appl
19	125.2	7.0	9551	2	US-08-800-644-93	Sequence 93, Appl
20	113.8	6.4	2338	1	US-08-425-069-1	Sequence 1, Appli
21	113.8	6.4	2338	2	US-08-317-744B-1	Sequence 1, Appli
22	112.6	6.3	2793	1	US-08-209-747-1	Sequence 1, Appli
23	112.6	6.3	2793	1	US-08-458-298-1	Sequence 1, Appli
24	103.2	5.8	1895	1	US-08-425-069-3	Sequence 3, Appli
25	103.2	5.8	1895	2	US-08-317-744B-3	Sequence 3, Appli
26	92.2	5.1	2830	2	US-09-010-928B-1	Sequence 1, Appli
C 27	89	5.0	1926	4	US-09-249-585A-4	Sequence 4, Appli

QY 1367 GCAAGAACGAGCGGAGCTGCTCGCGGAGAGGTGAACCTGCTGAGCAGGAGTGC 1426
DB 621 AGGAGGGGCGAGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG 680
QY 1427 AGGAGCTCGGGGCGGAGCGGCGCTTGGCCCGCGCATGCGGGCGCCACCTTCCCGAGG 1486
DB 681 AGGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGGGGCGAGG 740
QY 1487 AGCTCCCTGCTGAGCGGAGCTGAGCGGCTGCGGGCGGAGCTGCGGGGCGGAGCGGC 1546
DB 741 AGGAGGGGCGAGGAGGGGCGAGGAGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 800
QY 1547 AAGGCCATGACAGATGCTTCCGGCTTCCAGCATGAGCGCTGCTGTGGAAGAGGAGA 1606
DB 801 AGGCGCAGGAGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG 860
QY 1607 AGGAGAGGTGATTCAGTACCAAGAACAGCTGCGAGCAGAGTACGTCGCCCATGTACCAAGC 1666
DB 861 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGG 920
QY 1667 GGAACCGGCTGAGAGAGGGCTGCGAGGAGGGCTGCGAGGAGGGCTGCGAGGAGGGCTGCGAGGAGGG 1726
DB 921 GCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCTG 980
QY 1727 CCTTGGAGTTGACCTGGAAGGGGCTGACATCCCTACGAGG 1768
DB 981 AGGCCGGGCTGAGGAGGAGTGGAGGCGGGGCTGAGGAGG 1022

RESULT 2

US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 611411
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Pavan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 8.2%; Score 146; DB 3; Length 2580;

Best Local Similarity 49.5%; Pred. No. 2.5e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;
QY 1007 TGAGAGTTTACGAGGAGAACCGGAGCTCCGGCAGGAGCTCGAGAGCTTATGAGAGC 1066
DB 644 TGGAAACAGGAGCAGGAGCAGGAGCGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGG 703
QY 1067 AGGACCTGCTGGAGAACCAAGCTCAGTCTCCTACGAGGAGGAGAACAGCTTCCGCCCG 1126
DB 704 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763
QY 1127 CGTGTGAGGAGAGCCAGTGTGGAGGTGTCCAGAGTCCAGCGAGATCTCCTCTCTAAGC 1186
DB 764 AGGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 823
QY 1187 AGCAGCTGAAGAGTCTCCAGACGAGAGGTGAACCAAGGCTAGCGAGATCCTGGGTCTCA 1246
DB 824 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 883
QY 1247 AGGCACAGCTGAAGGACACCGCGGGCAAGCTGGAGGGCTTGGAGCTCAGGACCCAGGACC 1306
DB 884 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 943
QY 1307 TGGAGGGCGCCCTCGCGCACCAAGGGCTTGGAGCTTGGAGTCTGTGAGAAATGAGCTGAGC 1366
DB 944 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003
QY 1367 GCAAGAAAGCAGAGGCGGAGCTGCTGCGGAGAGAGGTGAACCTGCTGAGCAGGAGTGC 1426
DB 1004 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1063
QY 1427 AGGAGCTCGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1486
DB 1064 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123
QY 1487 ACGTCCCTGCTCGAGCGGAGCTGAGCGGCTTGGCGGCGGAGCTGCGGAGGAGGAGGAGG 1546
DB 1124 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183
QY 1547 AAGCCATGACAGATGCTTCCGCGCTTCCAGATGAGCGGCTGCTGAGAGGAGGAGGAGGAG 1606
DB 1184 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1243
QY 1607 AGGAGAGGTGATTTCAGTACCAAGAACAGCTGCGAGCAGAGTACGTGGCCATGTACAGC 1666
DB 1244 AGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1303
QY 1667 GGAACCGGCTTGGAGAGGCGCTTGCAGCAGCTTGGCACCTGGGAGCAGCGCCCGGAGC 1726
DB 1304 GGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1363
QY 1727 CCTTGGAGTTGACCTGGAAGGGCTGACATCCCTACGAGG 1768
DB 1364 AGGCCGGGCTGAGGAGGAGGAGTGGAGGCGGGGCTGAGGAGG 1405

RESULT 3

US-09-359-081-2
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Pavan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA

ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,081
FILING DATE: 22-Jul-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/050,863
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-8711
TELEFAX: (415) 949-8711
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match 8.2%; Score 146; DB 4; Length 2580;
Best Local Similarity 49.5%; Pred. No. 2.5e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 1007 TGCAGCTTCAGCAGGAGCGGCGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAGGAGC 1066
DB 644 TGGAAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 703
QY 1067 AGGACCTGTGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCCCG 1126
DB 704 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 763
QY 1127 CGCTGGAGAGACCCAGTGGAGGTGTGCGAGAGTCAAGGAGATCTCCCTCTGAAGC 1186
DB 764 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 823
QY 1187 AGCAGCTGAAGAGTCCAGCAGGAGGTGAAGCCAGGCTAGCAGATCTTGGGTCTCA 1246
DB 824 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 883
QY 1247 AGGCACAGCTGAAGACACGCGGGGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACC 1306
DB 884 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943
QY 1307 TGGAGGGCCCTCGCGCACCAAGGCGCTGGAGCTGGAGTCTGTGAGATGAGTGCAGC 1366
DB 944 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003
QY 1367 GCAAGAACAGCAGCGGAGTGTCTCGGAGAGAGTGAACCTGTGAGCAGGAGTGC 1426
DB 1004 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1063
QY 1427 AGGAGCTGGGGCCAGGCGCCCTGGCCCGGACATGGGGCCGCCACCTTCCCGGAGG 1486
DB 1064 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123
QY 1487 AGCTCCCTCGCTGAGCGGAGCTGGAGCGGCTCGCGGCCAGCTCGGGGAGGAGCGGC 1546
DB 1124 AGCAGGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183
QY 1547 AAGGCCATGACCATGTCTCTCGGCTTCACAGCATGAGGGCTGTGTGGAGGAGGAGA 1606
DB 1184 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1243

QY 1607 AGGAGAGGTGATTTCAGTACCAAGAAACAGCTGCAGCAGAGCTACGTGCGCCATGTACCA 1666
DB 1244 AGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1303
QY 1667 GGAACACAGCGCTGGAGAGAGCCCTGACAGAGCTGGGACAGCTGGGAGACAGCCCGGGAGC 1726
DB 1304 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1363
QY 1727 CTTGGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGG 1768
DB 1364 AGGCGGGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1405

RESULT 4

US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5978807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130.114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1

Query Match 8.2%; Score 146; DB 2; Length 5452;
Best Local Similarity 49.5%; Pred. No. 2.9e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 1007 TGCAGCTTCAGCAGGAGAGCGGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAGGAGC 1066
DB 2161 TGGAAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2102
QY 1067 AGGACCTGTCTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGAACCAAGCTTCGCGCCCG 1126
DB 2101 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2042
QY 1127 CGCTGGAGAGACCCAGTGGAGGTGTGCCAGAGTCAAGGAGATCTCCCTCTGAAGC 1186
DB 2041 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982
QY 1187 AGCAGCTGAAGAGTCCCAGACGAGGTGAACGCCAAGGCTAGCGAGATCTTGGGTCTCA 1246
DB 1981 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1922
QY 1247 AGGCACAGCTCAAGAGACACGCGGGGCAAGCTGGAGGGCTTGGAGCTGAGGACCCAGGACC 1306
DB 1921 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1862
QY 1307 TGGAGGGCCCTTCGCGCACCAAGGGCTTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGC 1366
DB 1861 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1802
QY 1367 GCAAGAACAGCAGGCGGAGCTGTGCGGAGAGGTGAACCTGTGAGCAGGAGGCTGC 1426
DB 1801 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1742
QY 1427 AGGAGCTCGCGGCCAGGCGCCCTTGGGCCCGGACATGGGGCCGCCACCTTCCCGGAGG 1486
DB 1741 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1682
QY 1487 ACCTCCCTGCTGACGCGGAGGTGGAGCGGCTTCGCGGCCCGAGCTGCGGAGGAGCGGC 1546

NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match 8.2%; Score 146; DB 4; Length 9600;
Best Local Similarity 49.5%; Pred. No. 3.2e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGGAGGCGAGCTCCGGCAGGAGCTCCAGAGCTCATGAAGGAGC 1066
Db 690 TGGACAGGAGCAGGAGCAGGAGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 749
Qy 1067 AGGACCTCTGGAGACCAAGCTCAGTCTCTAGAGAGGAGAGACAGCTTCGGCCCGC 1126
Db 750 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 809
Qy 1127 CGCTGGAGGAGACCCAGTGGAGGTGTCAGAGAGTCAAGGAGGAGATCTCCCTCTGAGC 1186
Db 810 AGGAGGCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 869
Qy 1187 AGCAGCTGAAGAGTCCAGCAGGAGGTGAACCCAGGCTAGCAGATCTCTGGTCTTCA 1246
Db 870 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 929
Qy 1247 AGGCACAGCTGAAGACACGCGGGGCAAGCTGGAGGGCTGAGCTGAGGAGCCAGGACC 1306
Db 930 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 989
Qy 1307 TGGAGGGCGCTTGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1366
Db 990 AGGAGGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1049
Qy 1367 GCAAGAGAACAGCAGGAGCTGTGCGGAGAGAGTGAACCTGTGAGCAGGAGGAGTGC 1426
Db 1050 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1109
Qy 1427 AGGAGCTCGGGCCAGCGCCCTTGGCCCGGACATGGGGCCGCCACCTTCCCGGAGG 1486
Db 1110 AGGAGGGGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1169
Qy 1487 ACCTCCCTGCTGAGCGGAGCTGGAGCGGCTGCGGCGCGAGCTGCGGAGGAGCGGC 1546
Db 1170 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1229
Qy 1547 AAGGCCATGACAGATGTCTCGGCTTCCAGCATGAGCGGTCTGTGGAAGAGGAGAGA 1606
Db 1230 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1289
Qy 1607 AGGAGAGGTGATTTCAGTACAGAAACAGCTGACAGAGTACGTGCGCCATGTACACAG 1666
Db 1290 AGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1349
Qy 1667 GGAAACACGGCTGGAGAGGCCCTGACAGAGCTGGACGTGGGACAGCGCCCGGGAGC 1726
Db 1350 GGCAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1409
Qy 1727 CTTTGGAGGTTGACCTGGAAGGGGCTGACATCCCTTACGAGG 1768
Db 1410 AGGCGGGGTGAGGAGGAGCAGTGGAGGCGGGGTGAGGAGG 1451

RESULT 7
US-07-884-811-15
; Sequence 15, Application US/07884811
; Patent No. 5316921
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/884,811
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-884-811-15

Query Match 8.2%; Score 146; DB 1; Length 10596;
Best Local Similarity 49.5%; Pred. No. 3.2e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGC 1066
Db 2224 TGGAAACAGGAGCAGGAGCAGGAGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 2283
Qy 1067 AGGACCTGTGGAGACCAAGCTCAGTCTACGAGAGGAGAGAACACAGCTTCGCGCCCG 1126
Db 2284 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2343
Qy 1127 CGCTGGAGGAGACCCAGTGGAGAGTGTGCCAGAGTCAAGGAGATCTCCCTCTCTGAAGC 1186
Db 2344 AGGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2403
Qy 1187 AGCAGCTGAAGGAGTCCAGACGAGGAGGTGAACCGCAAGGCTAGCGAGATCTCGGTCTCA 1246
Db 2404 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2463
Qy 1247 AGGCACAGCTGAAGGACACCGCGGGCAAGCTGGAGGGCTTGGAGCTGAGGAGCCAGGAGC 1306
Db 2464 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2523
Qy 1307 TGGAGGGCGCCCTCGCGCAACAAAGGGCTTGGAGCTGGAGGTCTGTGAGAAATGAGTGCAGC 1366
Db 2524 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2583
Qy 1367 GCAAGAGAACAGGCGGAGCTGTGCGGAGAGAGGTGAACCTGTCTGAGCAGGAGCTGC 1426


```

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-087-783A-15

Query Match      8.2%; Score 146; DB 1; Length 10596;
Best Local Similarity 49.5%; Pred. No. 3.2e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGGAGGCGGAGGCTCCGGCAGGAGCTCCAGAGCCCTCATGAAGGAGC 1066
Db 2224 TGGAAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2283

Qy 1067 AGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAGACACAGCTTCGCGCCCG 1126
Db 2284 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2343

Qy 1127 CGCTGGAGGAGACCCAGCTGGGAGGTGTGCCAGAAAGTCAGGCGAGATCTCCCTCTGAAGC 1186
Db 2344 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2403

Qy 1187 AGCAGCTGAAGGAGTCCACAGACGGAGGTGAACGCAAGGCTTAGCGAGATCTCGGTCTCA 1246
Db 2404 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2463

Qy 1247 AGGCACAGCTCAAGACACGCGGGGCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACC 1306
Db 2464 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2523

Qy 1307 TGGAGGGCGCCCTGCGCACCAAGGGCCCTGGAGCTGGAGGTGTGTGAGATGAGCTGCAGC 1366
Db 2524 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2583

Qy 1367 GCAGAAGAACGAGGCGGAGCTGTGCGGGAGAAAGGTGAACCTGTCGAGCAGGAGCTGC 1426
Db 2584 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2643

Qy 1427 AGGAGCTGCGGGCCAGCCCGCCCTTGGCCCGGACATGGGGCCGCGCCACCTTCCCGGAGG 1486
Db 2644 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2703

Qy 1487 ACCTCCCTGCCCCTGCAGCGGAGCTGGAGCGCTGCGGCGCCGAGCTCGGAGGAGCGGC 1546
Db 2704 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2763

Qy 1547 AAGGCCATGACAGATGTCTTCGCGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGA 1606
Db 2764 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2823

Qy 1607 AGGAGAGGTTATTCAGTACCAGAAACAGCTGCAGCAGAGTACGTGCGCATGTACCAGC 1666
Db 1607 AGGAGAGGTTATTCAGTACCAGAAACAGCTGCAGCAGAGTACGTGCGCATGTACCAGC 1666
```

```

Db 2824 AGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2883
Qy 1667 GGAACACGCGCTCGAGAAAGGCCCTCTGACAGAGCTGGCACTGGGGACACAGCCCGGGGAGC 1726
Db 2884 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2943
Qy 1727 CCTTGGAGGTTGACCTCGAAAGGGCTCACATCCCTACGAGG 1768
Db 2944 AGGCCGGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2985

RESULT 10
US-08-194-088B-15
; Sequence 15, Application US/08194088B
; Patent No. 5580963
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,088B
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallegos, R. Thomas
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: 755D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-2614
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-194-088B-15

Query Match      8.2%; Score 146; DB 1; Length 10596;
Best Local Similarity 49.5%; Pred. No. 3.2e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1066
Db 2224 TGGAAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2283

Qy 1067 AGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAGACACAGCTTCGCGCCCG 1126
Db 2284 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2343

Qy 1127 CGCTGGAGGAGACCCAGCTGGGAGGTGTGCCAGAAAGTCAGGCGAGATCTCCCTCTGAAGC 1186
Db 2344 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2403

Qy 1187 AGCAGCTGAAGGAGTCCACAGACGGAGGTGAACGCAAGGCTTAGCGAGATCTCGGTCTCA 1246
Db 2404 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2463

Qy 1247 AGGCACAGCTCAAGACACGCGGGGCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACC 1306
Db 2464 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2523

Qy 1307 TGGAGGGCGCCCTGCGCACCAAGGGCCCTGGAGCTGGAGGTGTGTGAGATGAGCTGCAGC 1366
Db 2524 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2583

Qy 1367 GCAGAAGAACGAGGCGGAGCTGTGCGGGAGAAAGGTGAACCTGTCGAGCAGGAGCTGC 1426
Db 2584 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2643

Qy 1427 AGGAGCTGCGGGCCAGCCCGCCCTTGGCCCGGACATGGGGCCGCGCCACCTTCCCGGAGG 1486
Db 2644 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2703

Qy 1487 ACCTCCCTGCCCCTGCAGCGGAGCTGGAGCGCTGCGGCGCCGAGCTCGGAGGAGCGGC 1546
Db 2704 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2763

Qy 1547 AAGGCCATGACAGATGTCTTCGCGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGA 1606
Db 2764 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2823

Qy 1607 AGGAGAGGTTATTCAGTACCAGAAACAGCTGCAGCAGAGTACGTGCGCATGTACCAGC 1666
Db 1607 AGGAGAGGTTATTCAGTACCAGAAACAGCTGCAGCAGAGTACGTGCGCATGTACCAGC 1666
```


Query Match	7.7%	Score 138.4;	DB 2;	Length 3489;
Best Local Similarity	47.0%	Pred. No. 7.3e-18;		
Matches 461;	Conservative	0;	Mismatches 516;	Indels 3;
				Gaps 1;

Qy	790	CAGGAGCTGGAGCAGAGACTGTTGGAGAGGGAGGGCGCCTCCAGAAAGTGTGCAGCGCAGC	849
Db	1639		
Qy	850	TTTTCAGAGAGAGGCTTTCCTCCAGCCTGGCCTACGAGAGCGCGCGCGCTGCAGG	909
Db	1699		
Qy	910	GACGAGCTGGAGGCGCGGAGCCGCAAGCTCAAGCAGGCGCTTCGACAGG	969
Db	1759	CGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGAGCCACAGCAGCGGGAGCCACAGCAG	1758
Qy	970	AGCCAGCGCGCAGCAGGCTTCGCACTTCGAGGTACTGCAGCTTCAGCAGGAGAGACGG	1029
Db	1819	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	1878
Qy	1030	CAGCTCCGCGAGGAGCTTCGAGAGCCTCATGAAGGAGCAGGACCTGCTGGAGAGCCAAAGCTC	1089
Db	1879	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG	1938
Qy	1090	AGGTCTTACGAGAGGGAGAAACAGCTTCGGCCCCCGCTGGAGGAGACCCAGTGGGAG	1149
Db	1939	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	1998
Qy	1150	GTGTCCAGAGTCAAGCGAGATCTCCCTCTCTGAAGCAGCAGCTGGAAGGAGTCCCAGACG	1209
Db	1999	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG	2058
Qy	1210	GAGGTGAACGCAAGGCTAGCGAGATCCTGGGTTCTCAAGGCAACAGCTGGAAGGACACGCGG	1269
Db	2059	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAG	2118
Qy	1270	GGCAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCTTCGCGCACCAAG	1329
Db	2119	GAGCAGCAGGATGAGCAGGAGACACAGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAG	2178
Qy	1330	GGCCTGGAGCTGGAGGCTCTGTGGAATGAGCTGCAGCGCAAGAGAACGAGGCGGAGCTG	1389
Db	2179	GATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGCAGGATGAG	2238
Qy	1390	CTGCGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGCGCCAGGCGGCC	1449
Db	2239	CAGCAGCAGCAGGATGAGCAGCAGCAGGAGCAGGAGCAGCAGGAGCAGGAGCAGGAG	2298
Qy	1450	CTGCGCGCGCATGCGGCGCCGCCACCTTCCCGGAGGACGCTCCCTGCGCTGCAGCGGAG	1509
Db	2299	CAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGATCAGGAGCAGGAGTTAGA--GGAG	2355
Qy	1510	CTGAGCGGCTGCGGCGGAGCTCGGGAGGAGCGGCAAGCCATGACCATGTCCTCG	1569
Db	2356	CAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG	2415
Qy	1570	GGCTTCCAGCATGAGCGGCTCGTGTGAAGGAGGAGAGAGGAGTGAATTCAGTACCAAG	1629
Db	2416	GAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA	2475
Qy	1630	AAACAGCTGCAGAGACTACGTGGCCATGTACAGCGGAAACCAAGCCCTGGAGAGAGGCC	1689
Db	2476	GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGTTGAA	2535
Qy	1690	CTGCAGCAGCTGGCAGCTGGGGACAGCGCCGGGAGCCCTTGGAGGTTGACCTCGAAGGG	1749
Db	2536	GAGCAGAGCAGGAGGTGAGAGCAAGAGCAGGAGCAGGAGCAGGAGCAGGAGTTAGAGAG	2595
Qy	1750	GCTGACATCCCTTACGAGGA	1769
Db	2596	GTGGAGGAGCAAGAGCAGGA	2615

```

; Patent No. 6322792
;
; GENERAL INFORMATION:
;
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestas, Mary E.
; APPLICANT: Kaye, Kenneth M.
;
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EPISOME PERSISTENCE
;
; FILE REFERENCE: 16412-10001R
;
; CURRENT APPLICATION NUMBER: US/09/2998,568
; CURRENT FILING DATE: 1999-04-21
;
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
;
; NUMBER OF SEQ ID NOS: 3
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
;
; LENGTH: 3489
;
; TYPE: DNA
;
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
;
; US-09-298-568-1

```

	Query Match	7.7%	Score 138.4	DB 4	Length 3489
	Best Local Similarity	47.0%	Pred. No. 7.3e-18		
	Matches 461	Conservative 0	Mismatches 516	Indels 3	Gaps 1
Qy	790	CAGGAGCTGAGCAGAAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGC	849		
Db	1639				
Qy	850	TTTGGAGAGAGAGCTTGCTTCAGCCTTGGCTTACGAGGAGCGCGCGCGCTGCAGG	909		
Db	1699	CGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGCGGAGCCACAGCAG	1758		
Qy	910	GACGAGCTGAGGGCCGGAGCCCAAGGGGGCAACAAGCTCAAGCAGGCGCTTCGAGAG	969		
Db	1759	CGGGAGCCACAGCAGCGGGAGCCACAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG	1818		
Qy	970	AGCCAGCGCGCAGCAGAGGTCCTGCACCTGCAGTACTGCAGCTTCAGCAGAGAAAGCGG	1029		
Db	1819	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	1878		
Qy	1030	CAGCTCCGGCAGGAGCTTCGAGAGCCTTCATGAAGGAGCAGGACCTGTGTGAGACCAAGTCT	1089		
Db	1879	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	1938		
Qy	1090	AGGTCCTACAGAGGGAGAAACACAGCTTCGGCCCGCGCTGGAGGAGACCCAGTGGGAG	1149		
Db	1939	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	1998		
Qy	1150	GTGTGCCAGAAATCAGGCGAGATCTCCCTCTGAAAGCAGCAGCTGAAAGAGTCCCCAGCG	1209		
Db	1999	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	2058		
Qy	1210	GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACAGTGAAGGACACGGCG	1269		
Db	2059	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGGAGCAGCAGGATGAGCAG	2118		
Qy	1270	GGCAGCTGAGGGCTTGGAGCTGAGGACCCAGGACCTGGAGGGCCCTTGGCCACCAAG	1329		
Db	2119	GAGCAGCAGGATGAGCAGGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAG	2178		
Qy	1330	GGCCTGAGAGCTGTGTGAGATGAGCTGCAGCGCAAGAAAGAACAGGCGCGAGCTG	1389		
Db	2179	GATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAG	2238		
Qy	1390	CTGCGGAGAGGTGAACCTGTGTGAGCAGGAGCCTGAGGAGCTGCGGGCCCAAGGCGGCC	1449		
Db	2239	CAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGGAGCAGGAGCAGCAGGAGCAGGAG	2298		
Qy	1450	CTGCGCCCGCACATGGGGCGGCCACCTTCCCGAGGACCTTCCCTGCGCTTCAGCGGGAG	1509		
Db	2299	CAGGATTAGAGGACAGGAGCAGGAGTTAGAGATCAGGAGCAGGAGGTTAG---GGAG	2355		
Qy	1510	CTGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGCGCAAGGCCATGACCAAGATGTCTCTCG	1569		

Db 2356 CAGGAGCAGGAGTTAGAGGACAGGAGCAGGAGTTAGAGGACGAGGAGTTAGAG 2415
Qy 1570 GGCTTCCAGCATGAGCGGCTCGTGTGGAAGAGGAGGAGGAGGAGGAGTTAGTTCAGTACCAG 1629
Db 2416 GAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGAGTTA 2475
Qy 1630 AAACAGCTGCAGCAGAGCTAGCTGCGCCATGTACAGCGGAAACAGCGCTCGAGAAAGGCC 1689
Db 2476 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGAGTGA 2535
Qy 1690 CTGAGCAGCTGGCAGCTGGGAGCAGCGCCGGGAGCCCTTGGAGGTTGACTTGAAGGG 1749
Db 2536 GAGCAAGAGCAGGAGGTTGAGAGAGCAAGAGCAGGAGCAGGAGGAGGAGGAGTTAGAGGAG 2595
Qy 1750 GCTGACATCCCTACGAGGA 1769
Db 2596 GTGGAGGAGCAGGAGCAGGA 2615

RESULT 15

US-08-770-379-20/c
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770.379
; FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52342
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 32207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-770-379-20

Query Match 7.7%; Score 138.4; DB 2; Length 32207;
Best Local Similarity 47.0%; Pred. No. 1.1e-17;
Matches 461; Conservative 0; Mismatches 516; Indels 3; Gaps 1;
Qy 790 CAGGAGCTGGAGCAGAGCTGTTCAGAGAGGAGGAGGCGCCCTCCAGAGCTCAGCGCAGC 849
Db 20358 CAGGAGCCACAGCAGCGGAGGCCACAGCAGCGGAGGCCCCAGCGGAGGCCACAGCAG 20299

Qy 850 TTTGAGGAGAGGAGCTTGGCTTCAGCTCGGCTTACGAGAGCGCGCGCGCTGTCAGG 909
Db 20298 CGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGCCACAGCAG 20239
Qy 910 GACGAGCTGGAGGCGCGGAGCCCAAGGCGGCAACAGCTCAGAGCAGGCTCCAGAG 969
Db 20238 CGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGTTAGCAGCAG 20179
Qy 970 AGCCAGCGCGCAGCAGAGGCTCTGCACCTGTCAGGTACTGTCAGCTTTCAGCAGGAGGAGCGG 1029
Db 20178 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 20119
Qy 1030 CAGTCCCGGAGGAGCTTCAGAGCTTCATGAGAGCAGGAGCAGCTGCTGAGAGCAAGCTC 1089
Db 20118 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 20059
Qy 1090 AGGTCTTACGAGGAGGAGAGCCAGCTTCGGCCCCCGCTGGAGGAGACCCAGTGGGAG 1149
Db 20058 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 19999
Qy 1150 GTGTCCAGAGTCAGGCGGAGATCTCCCTCTCTGAAGCAGCAGCTGAAAGGAGTCCACAGC 1209
Db 19998 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 19939
Qy 1210 GAGGTGAACGCCAAGGCTTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACCGCG 1269
Db 19938 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAG 19879
Qy 1270 GGCNAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTCGGCACCAAG 1329
Db 19878 GAGCAGCAGGATGAGCAGGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAGCAG 19819
Qy 1330 GGCCTGAGCTGGAGGCTGTGTGAGAACTGTGAGCGCGCAAGAGAAACAGCGCGGAGCTG 1389
Db 19818 GATCAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGCAGGATGAG 19759
Qy 1390 CTGCGGAGAGGTGAACCTGTGTGAGCAGGAGCTGCAGAGCTGCGGCGCCAGCGCCGCC 1449
Db 19758 CAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGCAGGAGCAGCAGGAGGAGCAGGAG 19699
Qy 1450 CTGCCCCCGACATGGGCGCGCCACCTTCCCCGAGGAGCTCCCTGCGCTGCAGCGGAG 1509
Db 19698 CAGGAGTTAGAGGAGCAGGAGCAGGATGAGGATCAGGAGCAGGAGTTAGA---GGAG 19642
Qy 1510 CTGAGCGCGCTGCGGCGCGGAGCTGCGGAGGAGCGGCAAGGCCATGACCAGATGTCTCTCG 1569
Db 19641 CAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGATGAGGAGCAGGAGCAGGAGTTAGAG 19582
Qy 1570 GGCTTCCAGCATGAGCGGCTCGTGTGGAAGAGGAGAGGAGAGGTTGATTTCAGTACAG 1629
Db 19581 GAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGAGTTA 19522
Qy 1630 AAACAGCTGCAGCAGCAGCTAGTGGCCATGTACAGCGGAAACAGCGCTCGGAGAGGCC 1689
Db 19521 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGAGTGA 19462
Qy 1690 CTGAGCAGCTGGCAGCTGGGAGCAGCGCCGGGAGCCCTTGGAGGTTGACTTGAAGGG 1749
Db 19461 GAGCAAGAGCAGGAGGTTGAGAGAGCAAGAGCAGGAGCAGGAGGAGGAGGAGTTAGAGGAG 19402
Qy 1750 GCTGACATCCCTACGAGGA 1769
Db 19401 GTGGAGGAGCAGGAGCAGGA 19382

Search completed: June 15, 2003, 03:45:50
Job time : 107.088 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 02:18:56 ; Search time 260.718 Seconds
(without alignments)
9947.427 Million cell updates/sec

Title: US-09-513-888C-3
 Perfect score: 1791
 Sequence: 1 atgqqcagcgtcactagcct.....tcatagccactgagatctga 1791

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	808.8	45.2	2411	9	US-09-764-904-89	Sequence 89, Appl
2	808.8	45.2	2411	9	US-10-091-548-89	Sequence 89, Appl
3	808.8	45.2	2411	9	US-10-074-095-1133	Sequence 1133, Ap
4	808.8	45.2	2411	10	US-09-764-860-1133	Sequence 1133, Ap
5	560	31.3	560	9	US-09-764-904-17	Sequence 17, Appl
6	560	31.3	560	9	US-10-091-548-17	Sequence 17, Appl
7	560	31.3	560	9	US-10-074-095-269	Sequence 269, App
8	560	31.3	560	10	US-09-764-860-269	Sequence 269, App
9	314	17.5	5257	9	US-10-171-581-63	Sequence 63, Appl
10	258.2	14.4	1617	10	US-09-925-300-60	Sequence 60, Appl
11	131.2	7.3	2108	10	US-09-962-832-225	Sequence 225, App
12	114.6	6.4	1852	10	US-09-969-852-4	Sequence 4, Appl
13	114	6.4	561	10	US-09-917-800A-194	Sequence 194, App
14	107.2	6.0	14800	10	US-09-954-456-1601	Sequence 1601, Ap
15	105.2	5.9	23585	9	US-09-764-891-6987	Sequence 6987, Ap
16	89	5.0	7065	9	US-09-991-496-115	Sequence 115, App
17	89	5.0	7065	10	US-09-874-923-115	Sequence 115, App
18	87.2	4.9	3773	10	US-09-925-302-47	Sequence 47, Appl
19	87.2	4.9	6457	10	US-09-880-107-3389	Sequence 3389, Ap

ALIGNMENTS

RESULT 1

US-09-764-904-89
; Sequence 89, Application US/09764904

; Sequence of: application
; Patent No. US20020171
; GENERAL INFORMATION:

Query Match 45.2%; Score 808.8; DB 9; Length 2411;

Qy	345	GGGCTCCGAGAGGGTGCAGTGAAGCCCAAGCCTTTCAGCCTGTGCTGCACCGTCAAG	404
Db	1306	GGGCTCCGAGAGGGTGCAGTGAAGCCCAAGCCTTTCAGCCTGTGCTGCACCGTCAAG	1365
Qy	405	AGGCATCTTGCACTCTCTCCCGAGAGTGCAGCCACCAAGCTGCACCCGCCCTCCAGA	464
Db	1366	AGGCATCTTGCACTCTCTCCCGAGAGTGCAGCCACCAAGCTGCACCCGCCCTCCAGA	1425
Qy	465	CAAGCCCAAGGACGAGAGCTGAAGCCTGCCTGTGCTCTGGGGCGCTGTCAAGATCCGG	524
Db	1426	CAAGCCCAAGGACGAGAGCTGAAGCCTGCCTGTGCTCTGGGGCGCTGTCAAGATCCGG	1485
Qy	525	CGGGAATCCATGTTCAGCCTGCCACACACAGCACCGACGAGCTACCAAGCTGGACCC	584
Db	1486	CGGGAATCCATGTTCAGCCTGCCACACACAGCACCGACGAGCTACCAAGCTGGACCC	1545
Qy	585	GCTGGTCAACCCGTGGGACCCCAAGCCGTTTTGGGGCTCCGCCCAACAATCAACCCA	644
Db	1546	GCTGGTCAACCCGTGGGACCCCAAGCCGTTTTGGGGCTCCGCCCAACAATCAACCCA	1605

QY 645 GGGCATCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTTCTCCGACGG 704
Db 1606 GGGCATCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTTCTCCGACGG 1665
QY 705 AGGTAGCAAGCTGGGCCACTCGAAACAAGGCGAGACAAGGGCCCTCGTGTGTCGCTCCCC 764
Db 1666 AGGTAGCAAGCTGGGCCACTCGAAACAAGGCGAGACAAGGGCCCTCGTGTGTCGCTCCCC 1725
QY 765 CATCTCCAGGACGAGTGCAGATCCAGAGCTGAGAGCTGAGAGAGAGAGGG 824
Db 1726 CATCTCCAGGACGAGTGCAGATCCAGAGCTGAGAGCTGAGAGAGAGAGGG 1785
QY 825 CGCCCTCCAGAGCTGCAGCGAGCTTTGAGAGAGAGAGCTTGCCTCCAGCTGGCCTA 884
Db 1786 CGCCCTCCAGAGCTGCAGCGAGCTTTGAGAGAGAGAGCTTGCCTCCAGCTGGCCTA 1845
QY 885 CGAGAGCGCGCGCGCGCTGCGAGGACGAGCTGAGGGCCCGGAGCCCAAGGCGGCAA 944
Db 1846 CGAGAGCGCGCGCGCGCTGCGAGGACGAGCTGAGGGCCCGGAGCCCAAGGCGGCAA 1905
QY 945 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCGAGAGCTTGCACCTGCAGGT 1004
Db 1906 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCGAGAGCTTGCACCTGCAGGT 1965
QY 1005 ACTGCAGCTTCAGCAGGAGAGCGCGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAAGGA 1064
Db 1966 ACTGCAGCTTCAGCAGGAGAGCGCGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAAGGA 2025
QY 1065 CGAGGACCTGTGGAGACCAAGCTCAGGTCCTTACGAGAGGAGAGACCAAGCTTCGGGCC 1124
Db 2026 CGAGGACCTGTGGAGACCAAGCTCAGGTCCTTACGAGAGGAGAGACCAAGCTTCGGGCC 2085
QY 1125 CGCGCTGAGAGACCAAGCTGAGAGGTGTGCC 1156
Db 2086 CGCGCTGAGAGACCAAGCTGAGAGGTGTGCC 2117

RESULT 2

US-10-091-548-89
; Sequence 89, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122CI
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-89

Query Match. 45.2%; Score 808.8; DB 9; Length 2411;
Best Local Similarity 99.8%; Pred. No. 5.5e-178;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 345 GGGCTCCGAGAGGCTGAGTGAGGCCACACAGCTTCAAGCTGTGTGCTGCCACGCTCAGG 404
Db 1306 GGGCTCCGAGAGGCTGAGTGAGGCCACACAGCTTCAAGCTGTGTGCTGCCACGCTCAGG 1365
QY 405 AGCCATCTTGCACTCTCTCCCGAGAGTGCCAGCACCAGCTGCAACCCCGCCCTCCAGA 464
Db 1366 AGCCATCTTGCACTCTCTCCCGAGAGTGCCAGCACCAGCTGCAACCCCGCCCTCCAGA 1425
QY 465 CNAAGCCAGGAGGAGGAGCTGAAGCTTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 524
Db 1426 CNAAGCCAGGAGGAGGAGCTGAAGCTTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 1485

QY 525 CCGGAATCTCATGTCCAGCTTGCCCAACACAGCACCAGCAGCAGCTTACAGCTGGACCC 584
Db 1486 CCGGAATCTCATGTCCAGCTTGCCCAACACAGCACCAGCAGCAGCTTACAGCTGGACCC 1545
QY 585 GCTGGTCAACCCGTGGGACCCCAAGCCGTTTGGGGGTTCGCCCCACAACTCACCACA 644
Db 1546 GCTGGTCAACCCGTGGGACCCCAAGCCGTTTGGGGGTTCGCCCCACAACTCACCACA 1605
QY 645 GGGCATCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTTCTCCGACGG 704
Db 1606 GGGCATCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTTCTCCGACGG 1665
QY 705 AGGTAGCAAGCTGGGCCACTCGAAACAAGGCGAGACAAGGGCCCTCGTGTGTCGCTCCCC 764
Db 1666 AGGTAGCAAGCTGGGCCACTCGAAACAAGGCGAGACAAGGGCCCTCGTGTGTCGCTCCCC 1725
QY 765 CATCTCCAGGACGAGTGCAGATCCAGGAGCTGAGAGCTGAGAGAGAGGG 824
Db 1726 CATCTCCAGGACGAGTGCAGATCCAGGAGCTGAGAGAGAGGG 1785
QY 825 CGCCCTCCAGAGCTGCAGCGAGCTTTGAGAGAGAGAGCTTGCCTCCAGCTGGCCTA 884
Db 1786 CGCCCTCCAGAGCTGCAGCGAGCTTTGAGAGAGAGAGCTTGCCTCCAGCTGGCCTA 1845
QY 885 CGAGAGCGCGCGCGCGCTGCGAGGACGAGCTGAGGGCCCGGAGCCCAAGGCGGCAA 944
Db 1846 CGAGAGCGCGCGCGCGCTGCGAGGACGAGCTGAGGGCCCGGAGCCCAAGGCGGCAA 1905
QY 945 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCGAGAGCTTCTGCACTGCAGGT 1004
Db 1906 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCGAGAGCTTCTGCACTGCAGGT 1965
QY 1005 ACTGCAGCTTCAGCAGGAGAGCGCGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAAGGA 1064
Db 1966 ACTGCAGCTTCAGCAGGAGAGCGCGAGCTCCGCGCAGGAGCTCGAGAGCTCATGAAGGA 2025
QY 1065 GCAGGACCTGTGGAGACCAAGCTCAGGTCCTTACGAGAGGAGAGACCAAGCTTCGGGCC 1124
Db 2026 GCAGGACCTGTGGAGACCAAGCTCAGGTCCTTACGAGAGGAGAGACCAAGCTTCGGGCC 2085
QY 1125 CGCGCTGAGAGACCAAGCTGAGAGGTGTGCC 1156
Db 2086 CGCGCTGAGAGACCAAGCTGAGAGGTGTGCC 2117

RESULT 3

US-10-074-095-1133
; Sequence 1133, Application US/10074095
; Publication No. US2003007704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08

```
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      45.2%; Score 808.8; DB 9; Length 2411;
Best Local Similarity 99.8%; Pred. No. 5.5e-178;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 345 GGGTCCGAGAGGGTGCAGTGAGGCCCCACAGCCTTCAAGCCTGTGTGTCGACCGGTGAGG 404
Db 1306 GGGTCCGAGAGGGTGCAGTGAGGCCCCACAGCCTTCAAGCCTGTGTGTCGACCGGTGAGG 1365

Qy 405 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCGCTCCAGCA 464
Db 1366 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCGCTCCAGCA 1425

Qy 465 CAAGCCCAAGAGAGAGAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 524
Db 1426 CAAGCCCAAGAGAGAGAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 1485

Qy 525 CCGGAATCTCATGTTCAGAGCTGCCACACAGCACCAGAGAGCTACAGCTGGAGCC 584
Db 1486 CCGGAATCTCATGTTCAGAGCTGCCACACAGCACCAGAGAGCTACAGCTGGAGCC 1545

Qy 585 GCTGGTCAACCCGTCGAGCCACCAAGCCGTTTGGGGGCTCCGCCACACATCACC 644
Db 1546 GCTGGTCAACCCGTCGAGCCACCAAGCCGTTTGGGGGCTCCGCCACACATCACC 1605

Qy 645 GGGCATCTGCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTGCTCTCCGACGG 704
Db 1606 GGGCATCTGCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTGCTCTCCGACGG 1665

Qy 705 AGGTAGCAAGCTGGGCGCACTCGAACAGCGAGACAAAGGCGCCCTCGTGTGTCGCTCC 764
Db 1666 AGGTAGCAAGCTGGGCGCACTCGAACAGCGAGACAAAGGCGCCCTCGTGTGTCGCTCC 1725

Qy 765 CATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 824
Db 1726 CATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 1785

Qy 825 CGCCCTCCAGAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCCTA 884
Db 1786 CGCCCTCCAGAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCCTA 1845

Qy 885 CGAGAGCGGCGCGCGCTGAGGAGAGCTGAGGAGCGCGGAGCCCAAGGCGGCA 944
Db 1846 CGAGAGCGGCGCGCGCTGAGGAGAGCTGAGGAGCGCGGAGCCCAAGGCGGCA 1905

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1133
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-860-1133

RESULT 4
US-09-764-860-1133
; Sequence 1133, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1133
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-860-1133

Query Match      45.2%; Score 808.8; DB 10; Length 2411;
Best Local Similarity 99.8%; Pred. No. 5.5e-178;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 945 CAAGCTCAAGCAGCGCTTCGCAGAGAGCCAGCGCGCAGCAGGTCTCTGCACTTCGAGGT 1004
Db 1906 CAAGCTCAAGCAGCGCTTCGCAGAGAGCCAGCGCGCAGCAGGTCTCTGCACTTCGAGGT 1965

Qy 1005 ACTGCAGCTTTCAGCAGGAGAGCGGCGCAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGGA 1064
Db 1966 ACTGCAGCTTTCAGCAGGAGAGCGGCGCAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGGA 2025

Qy 1065 GCAGGACCTCTGAGGACCAAGCTCAGGTCTTACGAGAGGAGAGACGAGCTTCGGGCC 1124
Db 2026 GCAGGACCTCTGAGGACCAAGCTCAGGTCTTACGAGAGGAGAGACGAGCTTCGGGCC 2085

Qy 1125 CGGCTGGAGAGACCCAGTGGAGGTGTGCC 1156
Db 2086 CGGCTGGAGAGACCCAGTGGAGGTGTGCC 2117

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1133
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-860-1133

Query Match      45.2%; Score 808.8; DB 10; Length 2411;
Best Local Similarity 99.8%; Pred. No. 5.5e-178;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 345 GGGTCCGAGAGGGTGCAGTGAGGCCCCACAGCCTTCAAGCCTGTGTGTCGACCGGTGAGG 404
Db 1306 GGGTCCGAGAGGGTGCAGTGAGGCCCCACAGCCTTCAAGCCTGTGTGTCGACCGGTGAGG 1365

Qy 405 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCGCTCCAGCA 464
Db 1366 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCGCTCCAGCA 1425

Qy 465 CAAGCCCAAGAGAGAGAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 524
Db 1426 CAAGCCCAAGAGAGAGAGCTGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGG 1485

Qy 525 CCGGAATCTCATGTTCAGAGCTGCCACACAGCACCAGAGAGCTACAGCTGGAGCC 584
Db 1486 CCGGAATCTCATGTTCAGAGCTGCCACACAGCACCAGAGAGCTACAGCTGGAGCC 1545

Qy 585 GCTGGTCAACCCGTCGAGCCACCAAGCCGTTTGGGGGCTCCGCCACACATCACC 644
Db 1546 GCTGGTCAACCCGTCGAGCCACCAAGCCGTTTGGGGGCTCCGCCACACATCACC 1605

Qy 645 GGGCATCTGCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTGCTCTCCGACGG 704
Db 1606 GGGCATCTGCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTGCTCTCCGACGG 1665

Qy 705 AGGTAGCAAGCTGGGCGCACTCGAACAGCGAGACAAAGGCGCCCTCGTGTGTCGCTCC 764
Db 1666 AGGTAGCAAGCTGGGCGCACTCGAACAGCGAGACAAAGGCGCCCTCGTGTGTCGCTCC 1725

Qy 765 CATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 824
Db 1726 CATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 1785

Qy 825 CGCCCTCCAGAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCCTA 884
Db 1786 CGCCCTCCAGAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGGCCTA 1845

Qy 885 CGAGAGCGGCGCGCGCTGAGGAGAGCTGAGGAGCGCGGAGCCCAAGGCGGCA 944
Db 1846 CGAGAGCGGCGCGCGCTGAGGAGAGCTGAGGAGCGCGGAGCCCAAGGCGGCA 1905
```

Db 1786 CGCCCTCAGAAAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCTCCAGCTCGGCCCTA 1845
QY 885 CGAGAGCGCGCGCGCTGCGAGGACGAGCTGAGGCGCGCGAGCCCAAGAGCGGCAA 944
Db 1846 CGAGAGCGCGCGCGCTGCGAGGACGAGCTGAGGCGCGCGAGCCCAAGAGCGGCAA 1905
QY 945 CAAGCTCAAGCAGGCGCTGCGAGAGAGCCAGCGCGCAGCAGGTCTGCACTGAGGT 1004
Db 1906 CAAGCTCAAGCAGGCGCTGCGAGAGAGCCAGCGCGCAGCAGGTCTGCACTGAGGT 1965
QY 1005 ACTGAGCTTCAGCAGGAGAGCGCGAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGA 1064
Db 1966 ACTGAGCTTCAGCAGGAGAGCGCGAGCTCCGCGAGGAGCTCGAGAGCTTCATGAAGA 2025
QY 1065 CGAGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCC 1124
Db 2026 GCAGGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCC 2085
QY 1125 CGCGTGGAGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCC 1156
Db 2086 CGCGTGGAGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGCGCC 2117

RESULT 5

US-09-764-904-17
; Sequence 17, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-17

Query Match 31.3%; Score 560; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGAGTGCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGAGTGCAGTGAGGC 60
QY 371 CCACAGCCTTCAAGCCTGTGCTGCCAGCGTCAGAGGCCATCTGCACCTCTCCCGGAGA 430
Db 61 CCACAGCCTTCAAGCCTGTGCTGCCAGCGTCAGAGGCCATCTGCACCTCTCCCGGAGA 120
QY 431 GTGCCAGCCACAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 180
QY 491 CTGGCCTGTGCTCTGGGGCGCTGTCCAGCTCCGAGGCCATCTGCACCTCTCCCGGAGA 550
Db 181 CTGGCCTGTGCTCTGGGGCGCTGTCCAGCTCCGAGGCCATCTGCACCTCTCCCGGAGA 240
QY 551 CACACAGCACCAGCAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 610
Db 241 CACACAGCACCAGCAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 300
QY 611 GCCGTTTGGGGGCTCCCGCCCAACATCATCCAGGGCATCGTCTCCAGGACAGCAACA 670
Db 301 GCCGTTTGGGGGCTCCCGCCCAACATCATCCAGGGCATCGTCTCCAGGACAGCAACA 360
QY 671 TGATGAGCCTGAAGGCTCTGCTCTCCGAGGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 730

Db 361 TGATGAGCCTGAAGGCTCTGCTCTCCGAGGAGTGAAGCTGGGCCACTCGAACA 420
QY 731 AGGAGACAAAGGCGCCCTCGTGTGTGCTCCCATCTCCACGAGCAGTGCAGCATCC 790
Db 421 AGGAGACAAAGGCGCCCTCGTGTGTGCTCCCATCTCCACGAGCAGTGCAGCATCC 480
QY 791 AGGAGCTGGAGCAAGAGCTTTGAGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 850
Db 481 AGGAGCTGGAGCAAGAGCTTTGAGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 540
QY 851 TTGAGGAGAAAGAGCTTGCC 870
Db 541 TTGAGGAGAAAGAGCTTGCC 560

RESULT 6

US-10-091-548-17
; Sequence 17, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-17

Query Match 31.3%; Score 560; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGAGTGCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGAGTGCAGTGAGGC 60
QY 371 CCACAGCCTTCAAGCCTGTGCTGCCAGCGTCAGAGGCCATCTGCACCTCTCCCGGAGA 430
Db 61 CCACAGCCTTCAAGCCTGTGCTGCCAGCGTCAGAGGCCATCTGCACCTCTCCCGGAGA 120
QY 431 GTGCCAGCCACAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 180
QY 491 CTGGCCTGTGCTCTGGGGCGCTGTCCAGCTCCGAGGCCATCTGCACCTCTCCCGGAGA 550
Db 181 CTGGCCTGTGCTCTGGGGCGCTGTCCAGCTCCGAGGCCATCTGCACCTCTCCCGGAGA 240
QY 551 CACACAGCACCAGCAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 610
Db 241 CACACAGCACCAGCAGCTGCACCCCGCTCCAGACAAAGCCCAAGAGCAGAGCTGAAGC 300
QY 611 GCCGTTTGGGGGCTCCCGCCCAACATCATCCAGGGCATCGTCTCCAGGACAGCAACA 670
Db 301 GCCGTTTGGGGGCTCCCGCCCAACATCATCCAGGGCATCGTCTCCAGGACAGCAACA 360
QY 671 TGATGAGCCTGAAGGCTCTGCTCTCCGAGGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 730
Db 361 TGATGAGCCTGAAGGCTCTGCTCTCCGAGGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 420
QY 731 AGGAGACAAAGGCGCCCTCGTGTGTGCTCCCATCTCCACGAGCAGTGCAGCATCC 790
Db 421 AGGAGACAAAGGCGCCCTCGTGTGTGCTCCCATCTCCACGAGCAGTGCAGCATCC 480
QY 791 AGGAGCTGGAGCAAGAGCTTTGAGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 850
Db 481 AGGAGCTGGAGCAAGAGCTTTGAGAGGAGGCGCCCTCCAGAAAGCTGCGAGCAGCT 540

Qy 851 TTGAGGAGAGGAGCTTCCC 870
Db 541 TTGAGGAGAGGAGCTTCCC 560

RESULT 7

US-10-074-095-269
; Sequence 269, Application US/10074095
; Publication No. US2003007704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 31.3%; Score 560; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGTCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGTCAGTGAGGC 60
Qy 371 CCACAGCCTTCAAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 430

Db 61 CCACAGCCTTCAAGCCTGTGTGCTGCCAGGTCAGGAGCCATCTTGCACCTCTCCCGGAGA 120
Qy 431 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAGAGCTGAAGC 180
Qy 491 CTGGCCTGTGCTCTGGGGCGGTGTGCAGACTCCGGCCGGAACCTCCATGTCCAGCCTGCCCA 550
Db 181 CTGGCCTGTGCTCTGGGGCGGTGTGCAGACTCCGGCCGGAACCTCCATGTCCAGCCTGCCCA 240
Qy 551 CACACAGCACCAGCAGCAGCTACAGCTGGACCGCTGGTCCACACCCCTGGGACCCCAAA 610
Db 241 CACACAGCACCAGCAGCAGCTACAGCTGGACCGCTGGTCCACACCCCTGGGACCCCAAA 300
Qy 611 GCCGTTTTGGGGGCTCCGCCCAACAACATCACCCAGGGCATCGTCTCCAGACACAGCAACA 670
Db 301 GCCGTTTTGGGGGCTCCGCCCAACAACATCACCCAGGGCATCGTCTCCAGACACAGCAACA 360
Qy 671 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTPAGCAAGCTGGGGCCTCACTGAACA 730
Db 361 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTPAGCAAGCTGGGGCCTCACTGAACA 420
Qy 731 AGCAGACAAGGGCCCTCGTGTGCTGCCCTCCCATCTCCACGACAGTGCAGCATCC 790
Db 421 AGCAGACAAGGGCCCTCGTGTGCTGCCCTCCCATCTCCACGACAGTGCAGCATCC 480
Qy 791 AGGAGCTGGAGCAGCAAGCTGTTGGAGAGGGAGGGCCCTCCAGAAGCTGCAGCGCAGCT 850
Db 481 AGGAGCTGGAGCAGCAAGCTGTTGGAGAGGGAGGGCCCTCCAGAAGCTGCAGCGCAGCT 540
Qy 851 TTGAGGAGAGGAGCTTGCC 870
Db 541 TTGAGGAGAGGAGCTTGCC 560

RESULT 8
US-09-764-860-269
; Sequence 269, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 269
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-269

Query Match 31.3%; Score 560; DB 10; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGTCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGTCAGTGAGGC 60
Qy 371 CCACAGCCTTCAAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 430
Db 61 CCACAGCCTTCAAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTTGCACCTCTCCCGGAGA 120
Qy 431 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAGAGCTGAAGC 180
Qy 491 CTGGCCTGTGCTCTGGGGCGGTGTGCAGACTCCGGCCGGAACCTCCATGTCCAGCCTGCCCA 550
Db 181 CTGGCCTGTGCTCTGGGGCGGTGTGCAGACTCCGGCCGGAACCTCCATGTCCAGCCTGCCCA 240

QY 551 CACAGACACAGGAGGAGCTACAGCTGGACCCGCTGGTACACCCCGTGGAGCCCAAA 610
DB 241 CACAGACACAGGAGGAGCTACAGCTGGACCCGCTGGTACACCCCGTGGAGCCCAAA 300
QY 611 GCCGTTTGGGGGCTCCGCCACACATCATCCAGGGGATCGTCTCCAGGACAGCAACA 670
DB 301 GCCGTTTGGGGGCTCCGCCACACATCATCCAGGGGATCGTCTCCAGGACAGCAACA 360
QY 671 TGATGAGCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACCTCGAACA 730
DB 361 TGATGAGCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACCTCGAACA 420
QY 731 AGGAGACAAAGGCCCTCTGTGTCTCCGATCCCTCCATCTCCACGAGAGTGCAGCATCC 790
DB 421 AGGAGACAAAGGCCCTCTGTGTCTCCGATCCCTCCATCTCCACGAGAGTGCAGCATCC 480
QY 791 AGGAGCTGAGCAGAGGCTGTGGAGAGGAGGGGCCCTCCAGAGCTGCAGCGCAGCT 850
DB 481 AGGAGCTGAGCAGAGGCTGTGGAGAGGAGGGGCCCTCCAGAGCTGCAGCGCAGCT 540
QY 851 TTGAGGAGAAGGAGCTTCCC 870
DB 541 TTGAGGAGAAGGAGCTTCCC 560

RESULT 9

US-10-171-581-63
; Sequence 63, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 63
; LENGTH: 5257
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB011124
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-63

Query Match 17.5%; Score 314; DB 9; Length 5257;
Best Local Similarity 58.2%; Pred. No. 1.8e-63;
Matches 619; Conservative 0; Mismatches 400; Indels 45; Gaps 2;
QY 710 GCAAGCTGGGCCACTCGAACAGGAGCAGACAGGGCCCTCGTGTCTCCGCTCCGCCATCT 769
DB 2292 GCACCTGGGCTCTCGGAGGGCGGAGGTGGAGGCTGCTTTCGGGCTGCTCACGCG 2351
QY 770 CCACGGAGAGTGACAGCTCCAGGAGCTGGAGCAGAGCTGTGGAGAGGGGGCGGCC 829
DB 2352 CTTCCCGGAGTCACTATCCAGGAGCTGGAGGAGCGGCTGTGGAGAGGAGGAGG 2411
QY 830 TCAGAGCTGAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTGGCTGAGGAGG 889
DB 2412 TGGCAGCTCTGCGGCGCAGCTTGGAGCAGAGCGGCTGTGGCCAGGTACTGGAGG 2471
QY 890 AGCGGCGCGGCGCTGCAGGAGCAGCTGGAGGGCCCGGAGGCCCAAGCGGCAACAAGC 949
DB 2472 ACGGGAGAGAGCGTGGGAGCGGAGCTGGCCAGCTGGGAGGCTGCAGCGGAGG 2531
QY 950 TCAAGCAGGCTTCGACAGAGAGCCAGCGCGCAGCAGAGTCTGACCTGACGAGTACG 1009
DB 2532 TACAGCAGGTGGCCGAGCTGCGCCAGCGCGCCAGCGGCGCTACAGCTGCAGGTGTTGC 2591

QY 1010 AGCTTCAGCAGGAGAACCGCAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAGCAGG 1069
DB 2592 GGCTGCAGCAGGACAAGAGCAGCTGCAGGAGGAGCGGCCCGCTGATCGGAGCGGG 2651
QY 1070 ACCTGCTGGAGACCAAGCTCAGGTCTTACAGAGGAGAGAACCAAGCTTCGGCCCCCGCGC 1129
DB 2652 AAGAGCTGGAGGACNAAGTGGCCGCTTCCAGAGGAGCAGGCCGACTTCTCTGCCCGGA 2711
QY 1130 TGGAGGAGACCAAGTGGAGGCTGTCCAGAGTCCAGAGTCCAGAGATCTCCCTCTCAAGCAGC 1189
DB 2712 TAGAGGAACTAAGTGGAGGTGTCCAGAGGCTGCGGAGATCTCCCTCTCAAGCAGC 2771
QY 1190 AGCTGAAGGAGTCCACAGCGAGGTGAACGCCAAGGCTAGGAGATCTCGGCTCTCAAGG 1249
DB 2772 AGCTGAAGGAGTCCACAGCGGCGATGTGTCGAGAAAGTTGAGTGAGATCGTGGGACTGCGCT 2831
QY 1250 CACAGCTGAAGGACACCGCGGGCAAGCTGGAGGCGCTGGAGCTCAGGACCCAGACCTGG 1309
DB 2832 CGAGCTTCGGGAGGCGGGCTTCGCTGCGGAGAGAGGAGCAGCTGCTCAGCTGCTC 2891
QY 1310 AGGCGGCTTCGCGCACCAAGGGCTTGAGGCTGGAGTCTGTGAGAAATGAGCTCAGCGCA 1369
DB 2892 GGGACTCTTCAGCAGCAAGCAGGCCAGCTGGAGCTGGGCGAAGGCGCAGCTGCTGCGC 2951
QY 1370 AGAAGAACGAGGCGGAGCTCTCGGGAGAGGTGAACCTGCTGGAGCAGCAGCT----- 1424
DB 2952 CTGCTTCAAGCGGCGCTGACCCCGTGGACCCGCGGAGCCACAGATGCTCTGGCCA 3011
QY 1425 -----GCAGGAGCTGCGGGGCCAGGCGCCCTGGCGCGCAGCATGCGGC 1468
DB 3012 CCTCGAGAGCAGCAGGCTAAGATGCGCGCTCAGGCGGGGTGCGCTCCGCTCT 3071
QY 1469 CGCCCACTTCCCGAGGAGC-----TCCCTGCCCTGCAGC 1504
DB 3072 TGGTTTCGTCGAGCGGAGCGGAGGCTGCGGGAGAGCGGAGCGCGGCCCTCGCGC 3131
QY 1505 GGGAGCTGGAGCGCTGCGGGCGAGCTGCGGGAGGAGCGGCAAGGCGCATGACAGATGT 1564
DB 3132 GGGAGGTGGGCGGCTGCAGGCGGAGCTGCGGCTGAGCGGCGGCGCGGAGCGCGCAGG 3191
QY 1565 CCTCGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGTGTTCAGT 1624
DB 3192 GTGCCAGCTTCGCGAGGAGCGCGCTGTGGCTGGAGGAGAGGAGTGTTCAGT 3251
QY 1625 ACCAGAAACAGCTGCAGCAGAGTACGTGGCCATGTACAGCGGAAACAGCGCTTGAGA 1684
DB 3252 ACCAGAACAGCTGCAGCTGAGCTACGTGGAGATGTACAGCGCAACAGCAGCTGGAGC 3311
QY 1685 AGGCCCTGCAGCAGCTGGGAGCAGCGCGGGAGCC 1728
DB 3312 CGAGGCTGCGGGAGCGGCGGCGCGGAGGGGTGCAAGCAGCC 3355

RESULT 10

US-09-925-300-60
; Sequence 60, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 60
; LENGTH: 1617

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1590)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1592)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1595)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1617)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-60

Query Match      14.4%; Score 258.2; DB 10; Length 1617;
Best Local Similarity 60.2%; Pred. No. 1.3e-50;
Matches 523; Conservative 0; Mismatches 303; Indels 43; Gaps 4;

Qy 962 CGCAGAGAGCAGCGCGCAGCAGGTCCTGCACCTGCAGTCTGCAGCTTTCAGCAGG 1021
Db 33 CACAGCGGCACAGCGCGCCACAGTGTCTGCAGCTGCAGTGTTCAGCTGCA-CAGG 91

Qy 1022 AGAAGCGCAGCTCCGGCAGGAGCTTCGAGAGCTTCATGAAGAGCAGGACCTGTCTGGAGA 1081
Db 92 AGAAGCGCAATTGCAGGACGACTTCGCACAGCTGCTGCAGAGCGCGCAACAGCTGGAGC 151

Qy 1082 CCAAGCTCAGGCTTACAGAGGGAGAGACCCAGCTTCGCGCCCGCTGCAGAGAGCC 1141
Db 152 GCGCTGCGCCACCTTTGAGCGGGACAGCGGGAGC-TCGGGCGGAGGCTTGAGGAGACA 210

Qy 1142 AGTGGAGGTGTCCAGAGTTCAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGT 1201
Db 211 AGTGGAGGTGTCCAGAAATCAGCGGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGT 270

Qy 1202 CCCAGACGGAGGTGAACCCAGGCTAGCGAGATCTTGGGTCTCAAGGCACAGCTGAAGG 1261
Db 271 CTCAGGACAGAGCTGTGTGAGAGGCGCAGAGCTGGTGGCTCTGCGGGTGGCGCTCGGG 330

Qy 1262 ACACGCGGGCAAGCTGAGGCGCTGAGCTGAGGACCCAGGACCTGAGGCGCCCTGC 1321
Db 331 AGGCGCGTGTACGCTGCGGGTTCAGTGAAGGCGGCTGCGGGGTCTACAGAGGCGCGCC 390

Qy 1322 GCACCAAGGCTGAGCTGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAGAACAGG 1381
Db 391 GAGCTCGGAGCTGAGCTGAGGCTGTTCAGGAGCTGCAGCGACACCGCCAGGAAG 450

Qy 1382 CGGAGCTGCTCGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGG--- 1438
Db 451 CTGAGCAGCTGCGGGAGAAAGCTGGGCAGTTGGATGCTGAGCGCGCGGACTCCGGGAGC 510

Qy 1439 -----CCGAGCGCCCTGCGCCCGCAGATGG 1465
Db 511 CCCCTGTGCCACCTGCCACCGCTGACCCATTCCTCTGGCAGAGAGTATGAGGCCAAG 570

Qy 1466 GCGCGCCACCTTCCCGCAGAGCTTCCTGCGCTGACGCGGAGCTGAGCGGCTGCGGG 1525
Db 571 TGCAGCGGCGAGCAGCGCGGGTGGGGCAGCTTGCGGGCCAGGTGAGGCGATTTGCGGG 630

Qy 1526 CCGAGCTCGGGAGAGCGCAAGGCCATGACCCAGATGCTCTCGGGCTTCAGCATGAGC 1585
Db 631 TGGAGCTCAGCGGAGCGCGCGGGGTGAGGACAGCGGACAGCTTTGAGGGGAGC 690

Qy 1586 GGCTGCTGGAAGAGAGAGAGAGAGAGTGTATTCAGTACAGAAAACAGCTGCAGCAGA 1645
Db 691 GGCTGGCTGGCAGCAGAGAGAGAGAGTGTATCCGCTACAGAAAGCAGCTGCAGCACA 750

Qy 1646 GCTAGTGGCCATGTACCAGCGGACAGCGCTTGGAGAGCCCTGCAGCAGCTGGCAC 1705
Db 751 ACTACATCCAGATGTACCGGCGCAACCGCGAGCTAGAGCAGAGAGCTGCAGCAGCTCAGCC 810
```

```
Qy 1706 GTGGGACAGCGCCGGGAGCCCTTGGAGGTGACCTG-----GAAGGGGTGACATCCC 1760
Db 811 TGGAGCTGGAGCGCCGGGAGCTCGCTGACCTGGGCTTGGCGGAGCAGCCCTGCTGCTG 870

Qy 1761 CTACGAGACATCATAGCCACTGAGATCT 1789
Db 871 CCTGGAGAGATCACTGCTACTGAGATCT 899

RESULT 11
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225

Query Match      7.3%; Score 131.2; DB 10; Length 2108;
Best Local Similarity 48.4%; Pred. No. 3.3e-21;
Matches 429; Conservative 0; Mismatches 448; Indels 9; Gaps 2;

Qy 860 AGGAGCTTGCTCCAGCTTCGAGGAGCGCGCGCTGCGAGGAGGAGGAGGAGGAGGAG 919
Db 499 AGCAGCAGGAGGGGCACTGAAGCACCTAGAGCAGCAGGAGGAGGAGGAGGAGGAGG 558

Qy 920 AGGCGCGGAGCCCAAGGGCGCAACAGCTCAGCAGGCGCTTCGAGAGGAGGAGGAGGAG 979
Db 559 AGCAGCAGGAGGGGCGAGCTGGAGTCTCCAGAGCAGCAGGAGGAGGAGGAGGAGGAG 618

Qy 980 CGCAGCAGGCTCTCCACTGTCAGCTTTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 1039
Db 619 AGCAGCAGGAGGGGCGAGCTGGAGTCTCCAGAGCAGCAGGAGGAGGAGGAGGAGGAG 678

Qy 1040 AGGAGCTCGAGAGCTCATGAAGAGCAGGAGCTGCTGGAGACCAAGCTCAGGCTCCTACG 1099
Db 679 AGCAGCAGGAGGGGCGAGCTGGAGTCTCCAGAGCAGCAGGAGGAGGAGGAGGAGGAG 738

Qy 1100 AGAGGAGAGACAGCTTTCGCGCCCGCTGAGGAGGAGACCCAGTGGAGAGGTGTGCCAGA 1159
Db 739 AGCAGCAGGAGGGGCGAGCTGGAGTCTCTGAGCAGCAGGAGGAGGAGGAGGAGGAGG 798

Qy 1160 AGTCAGGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGTCCAGACGCGAGGTGAACG 1219
Db 799 AGCAGCAGGAGGGGAGCAGCTGAAGCCTCGAGACACCGAGGAGGAGGAGGAGGAGGAG 858

Qy 1220 CCAAGGCTAGCGAGATCTTGGGTCTCAAGGCACAGCTGAAGGAGACACGCGGGCGAGCTGG 1279
Db 859 AGGAGCAGATGGGCGAGCTGAAGTACCTGGAACAGCAGGAGGAGGAGGAGGAGGAGGAG 918

Qy 1280 AGGCGCTTGAGCTGAGGAGC-----CCAGGACCTGGAGGGGCGCCCTTGGCGACCAAGGGGCC 1333
Db 919 ATCAGCAGGAGAGAGCAGCAGAGCTCCAGAGCAGCAGATGGGGCAGCTGAAGCACCTGG 978

Qy 1334 TGGAGCTGGAGGTCTGTGAGATGAGTGCAGGCCCAAGAACAGCAGCGGAGGAGGAGGAG 1393
Db 979 AGCAGCAGGAGGGGCGAGCCTTAAGCATCTGGAGAGCAGGAGGAGGAGGAGGAGGAGGAG 1038

Qy 1394 GGGAGAAAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGGAGTGCAGGAGGAGGAGGAGGAG 1453
```


RESULT 14
US-09-954-456-1601
; Sequence 1601, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Seq6
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1601

RESULT 15

US-09-764-891-6987/c
; Sequence 6987, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6987
; LENGTH: 22585
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9701)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-6987

Query Match 5.9%; Score 105.2; DB 9; Length 22585;
Best Local Similarity 47.7%; Pred No. 4,4e-15;
Matches 401; Conservative 0; Mismatches 434; Indels 6; Gaps 3;

QY 915 GCTGAGGGCCCGGAGCCCAAGCGGCAACAAGCTCAAGCAGGCTTCGACAGAGAGCCA 974
Db |||||
10432 GCAGAGGGGACTGGGCTGCAGGAGGAGCGGGGCTGCAGGAGAGCGGGGCTGCAGGA 10373
QY 975 GCGCGCGCAGCAGGTCCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGAGAGCGCAGCT 1034
Db |||||
10372 GGGGACGGCGCTGCAGAGGAGATGCAGATGCAGAGGGGACGCGCAGATGCAGGAGGGGAC 10313
QY 1035 CCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGACCTGTGGAGACCAAGCTCAGGTC 1094
Db |||||
10312 GCAGTGCAGGAGGAGCGCAGATGCAGAGGGGACGGGGCTGCAGGAGGAGATGCAGGA 10253
QY 1095 CTACGAGAGGAGAGACAGCTTCGGCCCGCGCTGGAGGAGACCCAGTGGGAGGTGTG 1154
Db |||||
10252 GGGGCTGCAGGTACAGGAGGGGTTGGGGC---TGCAGGAGGGGACACAGATGCAGGAGGG 10196
QY 1155 CCAGAGTTCAGGCGAGATCTCCTCCTGAGCAGCAGCTGAGGAGTCCACAGCGAGGT 1214
Db |||||
10195 GCGCGAGATGCAGAGGGGACGGGGCTGCAGAGGAGATGCAGATGCAGGAGGGGATGCA 10136
QY 1215 GAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAA 1274
Db |||||
10135 GATGCAGGAGGAGATGCAGATGCAGAGGGGATGCAGGTGCAGGAGGGGACCGGGGCTGCA 10076
QY 1275 GGTGAGGGCCTGGAGCTGAGACCCAGACCTTGAGGGCGCCCTTGCAGCAACAGGGCCT 1334
Db |||||
10075 GGAGGGAGCGGGGCTGGAGGGGACGGGGCTACAGGAGGAGATGCAGATGCAGGAGGGGT 10016
QY 1335 GGAGCTGGAGGCTCTGTGAGATGAGCTGACGGCCAGAGACAGGCGGAGCTGCTGGG 1394
Db |||||
10015 GCACGTACAGGACTGGTTCGGGGCTGCAGAGGGGACGCGAGATGCAGGAGGGGCTCGGGCT 9956
QY 1395 GGAGAGGTGAACCTGTGAGCAGGAGCTGCAGGAGCTGCGGGGCCAGAGGCCCTGGC 1454
Db |||||
9955 GCAGGAGGGGAGGACATCCAGGAGGGGACCGGGTGCAGGAGGGGACCGGGTGCAGGA 9896
QY 1455 CCGGACATATGGGGCCGCCACCTTCCCGAGAGCTCCCTGCGCTGCAGCGGGAGCTGGA 1514
Db |||||
9895 GGGGGCGCAGGGGCGAGGAGGAGCGGGTGCAGGAGGGGACACAGGTTCAGGAGGGGAC 9836
QY 1515 GCGGCTGCGGGCCGAGCTGCGGAGGAGCGGACAGCCATGACAGATGCTCTCGGGCTT 1574
Db |||||
9835 GGGGCTGCAGGAGGAGATGCAGATCAGGGGGGTG-CAGGTACAGGAGGGGCTCGGGCT 9777
QY 1575 CCAGCATGAGCGGCTCGTGTGAGAGGAGGAGAGAGGTGATTTCAGTACCAAGACA 1634

Db 9776 GCAGAGGGGAGCGCAGATGCAGGAGGGGACGCGAGATGCAGAGGGGACACAGATGCAGAA 9717
QY 1635 GCTGCAGCAGAGCTACGTGGCCATGTACACAGCGGAACAGCGCTTGAGAGAGGCCCTGCA 1694
Db |||||
9716 GGGGGCGCAGATGCANGAAGCGACGCGGGGCTGCAGAGGAGATGCAGATGCAGAGGAGAT 9657
QY 1695 GCAGCTGCACGTGGG--GACAGCGCCGCGGAGCCCTTTGGAGGTTGACCTTGAAGGGGCT 1752
Db |||||
9656 GCAGATCCCTGGAGGGATGCAGGTGCAGAGGGGACGGGGCTGCAGGACGGGATGGGGCT 9597
QY 1753 G 1753
Db 9596 G 9596

Search completed: June 15, 2003, 06:49:17
Job time : 263.718 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 2493.59 Seconds
(without alignments)
11632.279 Million cell updates/sec

Title: US-09-513-888C-3
Perfect score: 1791
Sequence: 1 atggcgagcgtcagtagcct.....tcatagcactgagatctga 1791

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_othr.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	608.6	34.0	668	12 BF058214	BF058214 7k28d05.x
C 2	606.2	33.8	817	9 A1984777	A1984777 wr85b07.x
C 3	545	30.4	545	10 AW007737	AW007737 wr68e06.x
C 4	523.4	29.2	740	14 BQ769435	BQ769435 UI-M-F10-
C 5	489.6	27.3	547	12 BF590813	BF590813 7h42e07.x
C 6	447	25.0	447	9 A1042490	A1042490 ox62e04.x

7	412.2	23.0	596	10	AV663618
8	366.6	20.5	397	10	AW028197
C 9	338.2	18.9	344	10	AW016544
10	299.4	16.7	394	12	BF549120
11	298.4	16.7	394	12	BF554804
C 12	288.4	16.0	743	10	BE410921
C 13	286.8	16.0	673	10	BE384131
C 14	286	16.0	472	12	BF706253
C 15	272.4	15.2	658	17	AZ858490
C 16	271.2	15.1	725	9	AJ454524
C 17	256.4	14.3	407	10	AV663617
C 18	253	14.1	923	12	BG750395
19	237.2	13.2	932	14	BQ879505
20	236	13.2	635	12	BF037137
21	233.2	13.0	921	14	BQ652556
22	229.8	12.8	663	9	AL635635
23	229.8	12.8	991	14	BQ720762
24	229.2	12.8	907	14	BQ652760
25	228	12.7	846	12	BF705082
26	221	12.3	735	12	BE740090
27	218.4	12.2	806	14	BM947278
28	213.4	11.9	661	12	BF312957
29	213	11.9	700	13	BI854892
C 30	207.6	11.6	312	17	AZ595691
31	203.4	11.4	1115	12	BF784037
32	203	11.3	906	13	BG917774
33	200.4	11.2	894	13	BI520208
34	199.6	11.1	487	13	BI338021
35	199.4	11.1	1091	12	BG255093
36	198.6	11.1	1101	14	BQ051027
37	197.8	11.0	602	10	BE304332
38	197.6	11.0	723	10	BE312971
39	192.6	10.8	941	14	BQ955592
40	189.2	10.6	839	13	BI752956
41	185.4	10.4	460	12	BE757942
42	184.8	10.3	973	9	AL551801
43	183.4	10.2	926	12	BF583930
44	181	10.1	489	10	BE262105
C 45	174	9.7	531	10	BE276168

ALIGNMENTS

RESULT 1
BF058214/c
LOCUS
DEFINITION
7k28d05.x1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:3476769 3', similar to TR:Q9Y5W1 Q9Y5W1 FEZ1. [5] TR:Q9Y5W0 TR:Q9Y5V9 TR:Q9Y5V8 TR:Q9Y5V7 ;contains MER22.t3 TAR1 repetitive element ;, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BF058214
BF058214.1
EST.
human.
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 668)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

AV663618 AV663618
AW028197 wv84a01.x
AW016544 UI-H-BIOP
BF549120 UI-R-AO-a
BF554804 UI-R-E0-C
BE410921 601303579
BE384131 601272956
BF706253 280459 MA
AZ858490 2M0163110
AJ454524 AJ454524
AV663617 AV663617
BG750395 AGNCOURT
BQ879505 AGNCOURT
BF037137 601456991
BQ652556 AGNCOURT
AL635635 AL635635
BQ720762 AGNCOURT
BQ652760 AGNCOURT
BF705082 602274088
BE740090 601595016
BM947278 UI-M-EHOP
BF312957 601896225
BI854892 603381873
AZ595691 IM0408A04
BF784037 602107882
BG917774 602818743
BI520208 603071456
BI338021 361664 MA
BG255093 602369606
BQ051027 AGNCOURT
BE304332 601105731
BE312971 601150210
BQ955592 AGNCOURT
BI752956 603025713
BE757942 212520 MA
AL551801 AL551801
BF583930 602096941
BE262105 601146482
BE276168 601144424

Seg primer: -40UP from Gibco
High quality sequence stop: 441.
Location/Qualifiers
1. .668

FEATURES source

AI984777
LOCUS AI984777.1
DEFINITION w:85b07.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2494453 3', similar to TR:O60295 O60299 KIAA0552 PROTEIN. ; mRNA sequence.
ACCESSION AI984777
VERSION AI984777.1
KEYWORDS GI:5812054
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 817)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1307 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 428.
Location/Qualifiers
1. .817

BASE COUNT 105 a 185 c 226 g 150 t 2 others
ORIGIN

Query Match 34.0%; Score 608.6; DB 12; Length 668;
Best Local Similarity 95.2%; Pred. No. 3.9e-101;
Matches 637; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
QY 190 GACTCTTCTACATCAAGTACGAGCAGAGAGCCCGGGCTCCATCACCAGATTACAGC 249
DB 668 GACTCTTCTTAACATCAATGTAGCCAGTAAGCCGAGGCTCCATCCCAAGATTATCG 609
QY 250 GCATGTGCCAGGGGATTAGGGGGCCAGGCTGGGTGGATTGACCCGTCCACACCC 309
DB 608 GCATGTGCCAGGGAATAGAGGGCCAGATGGGGTGAATTTGACNCNTCCACACA 549
QY 310 CCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGTGCAGTGAGG 369
DB 548 CCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGTGCAGTGAGG 490
QY 370 CCCAAGCTTCAAGCCTGTGTGTCAGGCTCAGAGGCATCTGTCATCTCTCCCGGAG 429
DB 489 CCCAAGCTTCAAGCCTGTGTGTCAGGCTCAGAGGCATCTGTCATCTCTCCCGGAG 430
QY 430 AGTGCCAGCCACAGCTGCACCCGCCCTCCAGAACGCCCAAGGAGGAGCTGAAG 489
DB 429 AGTGCCAGCCACAGCTGCACCCGCCCTCCAGAACGCCCAAGGAGGAGCTGAAG 370
QY 490 CTTGCCCTGTGCTTGGGGCGCTGTGACACTCCGGCCGGAATCTCATGTCCAGCTGCC 549
DB 369 CTTGCCCTGTGCTTGGGGCGCTGTGACACTCCGGCCGGAATCTCATGTCCAGCTGCC 310
QY 550 ACACAGACACAGCAGCAGCTACAGCTGGACCGCTGCTCAGACCCGTGGGACCCACA 609
DB 309 ACACAGACACAGCAGCAGCTACAGCTGGACCGCTGCTCAGACCCGTGGGACCCACA 250
QY 610 AGCCGCTTTGGGGCTCCGCCCAACAATCAGAGGCGATCGTCTCCAGAGCAGCAAC 669
DB 249 AGCCGCTTTGGGGCTCCGCCCAACAATCAGAGGCGATCGTCTCCAGAGCAGCAAC 190
QY 670 ATGATGAGCCTGAAGGCTCTGTCTTCTCCAGAGGAGGTAGCAAGCTGGGCGACTCGAAC 729
DB 189 ATGATGAGCCTGAAGGCTCTGTCTTCTCCAGAGGAGGTAGCAAGCTGGGCGACTCGAAC 130
QY 730 AAGGAGACAGAGGGCCCTCTGTGTGCTGCTCCCTCCCTCCAGAGGAGGTAGCAAGCT 789
DB 129 AAGGAGACAGAGGGCCCTCTGTGTGCTGCTCCCTCCCTCCAGAGGAGGTAGCAAGCT 70
QY 790 CAGGAGCTGGAGCAGAACTGTGTGAGAGGGAGGGCGCCCTCCAGAACTGCAGGCGCAGC 849
DB 69 CAGGAGCTGGAGCAGAACTGTGTGAGAGGGAGGGCGCCCTCCAGAACTGCAGGCGCAGC 10
QY 850 TTTGAGGAG 858
DB 9 TTTGAGGAG 1

RESULT 2 AI984777/c

LOCUS AI984777.1
DEFINITION w:85b07.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2494453 3', similar to TR:O60295 O60299 KIAA0552 PROTEIN. ; mRNA sequence.
ACCESSION AI984777
VERSION AI984777.1
KEYWORDS GI:5812054
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 817)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1307 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 428.
Location/Qualifiers
1. .817

FEATURES source

BASE COUNT 133 a 235 c 266 g 180 t 3 others
ORIGIN

Query Match 33.8%; Score 606.2; DB 9; Length 817;
Best Local Similarity 94.0%; Pred. No. 1.1e-100;
Matches 717; Conservative 0; Mismatches 35; Indels 11; Gaps 8;
QY 112 TCCACAGGGCTGTGAGGTTGGCTTCTCCAGAGCTCCGGTACGCGCAAGTCCAGCTCC 171
DB 756 TCCACAGGGCTGTGTAAGTTT--GCTTTTCCCGGACTCC-GTCACGCAAGTCCAGCTCC 700
QY 172 AAAATGGGCAAGAGGAGACTTCTTACATCAAGGTACGCGAGAAAGCCCGGGGCTCC 231
DB 699 AAA--TGCAAGAGCGAGAGCTTCTTACATCAAGG-CAGCGAGAAAGCCCGGGGCTCC 643
QY 232 CATCACCAGATTACCGGCACTGTCCAGCGGGGATTAGGGGCCAGGCT-GGGGTGA 290
DB 642 CATCACCAGATAACAC-GCACGGTCCAGCGGGGATTAGGGGCCAGGTTGGGGGTGA 584
QY 291 CTTTGACCCCTCCACACCCCAAGCTCATGCCCTTCTCCAAATCAGTAGAAATGGGCTC 350
DB 583 CTNTGCCCGCTCCCAACCCCAAGGCTCATGCCCTTCTCCAAATCAGTAGAAATGGGCTC 524
QY 351 CGAAGAGGGTGCAGTGAGGGCCACAGGCTTTCAGGCTTGTGCTGCCAGGTCAGGACCAT 410

Db 523 GTAGAAGGTCAGTAAGGCCACACAGCCTTCAAGCTTGTGTGCACCGTTCAGGAGCCAT 464
Qy 411 CCTGCACCTCTCCCGGAGAGTGCAGCACCACAGCTGCACCCCGCTTCAGACAAAGCC 470
Db 463 CCTGCACCTCTCCCGGAGAGTGCAGCACCACAGCTGCACCCCGCTTCAGACAAAGCC 404
Qy 471 CAAGGAGCAGAGCTGAAGCCTGGCCCTGTGTCTCTGGGGCGTGTGCAGACTCCGGCCGGAA 530
Db 403 CAAGGAGCAGAGCTGAAGCCTGGCCCTGTGTCTCTGGGGCGTGTGCAGACTCCGGCCGGAA 344
Qy 531 CTCATGTTCAGCTTGCACACACAG-CACAGCAGCAGCTACAGCT--GGACCCGCT 587
Db 343 CTCATGTTCAGCTTGCACACACAGCAGCCCGCAGCAGAAAGTTACCAAGTTGGGACCCGCT 284
Qy 588 GGTACACCCGTGGGACCCACAAAGCCGTTTTTGGGGGCTCCGCCCAACATCACCACAGG 647
Db 283 GGTACACCCGTGGGACCCACAAAGCCGTTTTTGGGGGCTCCGCCCAACATCACCACAGG 224
Qy 648 CATGCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTCTCCGACGAGG 707
Db 223 CATGCTCTCCAGGACAGCAACATGATGAGCTGAAGCTCTGTCTCTCCGACGAGG 164
Qy 708 TAGCAAGCTGGCCACTCGAACAGGACAGCAAGGCGCCCTCGTGTCTCCGCTCCCCCAT 767
Db 163 TAGCAAGCTGGCCACTCGAACAGGACAGCAAGGCGCCCTCGTGTCTCCGCTCCCCCAT 104
Qy 768 TTCACGACAGTGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTTGGAGAGGGAGGCGC 827
Db 103 CTCACGACAGTGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTTGGAGAGGGAGGCGC 44
Qy 828 CCTCAGAAGCTGCAGCGAGCTTTGAGGAGAGGAGCTTGCC 870
Db 43 CCTCAGAAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCC 1

RESULT 3
AW007737/c
LOCUS
DEFINITION
IMAGE:2512642.3, similar to contains MER22.t3 TARI repetitive
element ;, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW007737 545 bp mRNA linear EST 09-MAR-2000
wt68e06.xl Soares thymus NHFth Homo sapiens cDNA clone
IMAGE:2512642.3, similar to contains MER22.t3 TARI repetitive
element ;, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

BASE COUNT 80 a 159 c 189 g 117 t
ORIGIN
Query Match 30.4%; Score 545; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;
Matches 545; Conservative 0; Mismatches 0; Indels -0; Gaps 0;
Qy 326 TCTCAATCAGTCTAGAAATGGGCTCCGAGAAAGGTGAGTGAGGCCACACAGCTTCAAGC 385.
Db 545 TCTCAATCAGTCTAGAAATGGGCTCCGAGAAAGGTGAGTGAGGCCACACAGCTTCAAGC 486
Qy 386 CTGTGCTGCCAGCTCAGGAGCCATCTCTGACTCTCTCCCGGAGAGTCCAGCCACAGC 445
Db 485 CTGTGCTGCCAGCTCAGGAGCCATCTCTGACTCTCTCCCGGAGAGTCCAGCCACAGC 426
Qy 446 TGCACCCCGCCCTCCAGACAAAGCCCAAGGAGCAGGAGCTCAAGCTCGCTGCTCTG 505
Db 425 TGCACCCCGCCCTCCAGACAAAGCCCAAGGAGCAGGAGCTCAAGCTCGCTGCTCTG 366
Qy 506 GGGCGCTGTCAAGACTCCGGCCGGAATCCATGTCTCAGCTCCCAACACACAGCAGCAGCA 565
Db 365 GGGCGCTGTCAAGACTCCGGCCGGAATCCATGTCTCAGCTCCCAACACACAGCAGCAGCA 306
Qy 566 GCAGCTACAGCTGGACCCGCTGTGTCAACCCGTGGAGCCACAGCCGCTTTGGGGGCT 625
Db 305 GCAGCTACAGCTGGACCCGCTGTGTCAACCCGTGGAGCCACAGCCGCTTTGGGGGCT 246
Qy 626 CCGCCCAACATCACCAGGAGCTCGTCTCCAGGACAGCAATGATGAGCTGAAGG 685
Db 245 CCGCCCAACATCACCAGGAGCTCGTCTCCAGGACAGCAATGATGAGCTGAAGG 186
Qy 686 CTCTGTCTCTCCGAGGAGTGAAGCTGGGCTCACTCGAAACAGGACAGCAAGGGCC 745
Db 185 CTCTGTCTCTCCGAGGAGTGAAGCTGGGCTCACTCGAAACAGGACAGCAAGGGCC 126
Qy 746 CCTCGTGTCTCCGCTCCCACTCTCCAGGACAGTGCAGATCCAGAGCTGGAGCAGA 805
Db 125 CCTCGTGTCTCCGCTCCCACTCTCCAGGACAGTGCAGATCCAGAGCTGGAGCAGA 66
Qy 806 AGCTGTTGGAGGAGGAGGCGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGGAGGAGC 865
Db 65 AGCTGTTGGAGGAGGAGGCGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGGAGGAGC 6
Qy 866 TTGCC 870
Db 5 TTGCC 1

RESULT 4
BQ769435
LOCUS
DEFINITION
IMAGE:6400797.5, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BQ769435 740 bp mRNA linear EST 26-JUL-2002
UI-M-FIO-byg-a-22-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone
IMAGE:6400797.5, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 740)
NIT-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa

Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccocao.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

Location/Qualifiers
1..596
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1BR037F07"
/clone_lib="Bos taurus brain fetus"
/tissue_type="brain"
/dev_stages="fetus"
/lab_host="DH10B"
/note="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"
114 a 196 c 170 g 103 t 13 others

Query Match 23.0%; Score 412.2; DB 10; Length 596;
Best Local Similarity 87.9%; Pred. No. 2.3e-65;
Matches 471; Conservative 0; Mismatches 61; Indels 4; Gaps 2;
QY 92 TCACGAAGCTCAACGGTATTCGACGGCTGCTGAGGTTGGCTTCCCGAGGACTCG 151
DB 61 TCACGAAGCTCAACGGTATTCGACGGCTGCTGAGGTTGGCTTCCCGAGGACTCG 120
QY 152 GTCACGGAAGTCCAGCTCCAAATAGGCAAGGCGAAGACTTCTTACATCAAGGTCA 211
DB 121 GACATGGCAAGTCCAGCTCCAAATAGGCAAGGCGAAGACTTCTTACATCAAGGTCA 180
QY 212 GCCAAGAACCCGGGCTCCCATCCACGCAATTACACGCACTGCCAGCGGGATTTAG 271
DB 181 GCCAAGAACCCGGGCTCCCATCCACGCAATTACACGCACTGCCAGTGGGNNNTAG 240
QY 272 GGGGCCAGGCTGGGTGACCTTTGACCGGTCCACACCCCGGAGCTCATGCCCTTCTCA 331
DB 241 GGGGCCAGGCTGGGTGACCTTTGACCGGTCCACACCCCGGAGCTCATGCCCTTCTCA 300
QY 332 ATCAGCTAGAAATGGCTCCGAGAAGGTGACGAGGCGCCACAGCTTCAAGCTGTGC 391
DB 301 ATCAGCTAGAGATGGTCTGAGAAGGTGCTGTGAGACCCACGCGCTTCAAGCGGTGC 360
QY 392 TGCCACGCTCAGAGCCATCTTGCATCTCTCCCGAGAGTGCCAGCCACAGCTGCACC 451
DB 361 TGCCACGCTCAGAGCCATCTTGCATCTCTCCCGAGAGTGCCAGCCACAGCTGCATC 420
QY 452 CCGCCCTCCAGACCCAGGAGGAGGAGTCAAGCTGCTGCTGTCTGGGGCGC 511
DB 421 CTGCGCTCTTGACAAAGCCAGGAGGAGGAGTCAAGCTGCTGCTGTCTGGGGCGC 480
QY 512 TGTCAGACTCCGGCGGAACTCCATGTCCAGCTGCGCCACACACAGCAGCAGCT 571
DB 481 TGTCAGACTCCGGCGGAACTCCATGTCCAGCTGCGCCACACACAGCAGCAGCT 540
QY 572 A-CCAGCTGACCCGCTGG---TCACACCGGTGGGACCCACAAAGCGCTTTGGGGG 623
DB 541 ACCCAGCTGGACCCGCTGGTGTCACTCTGTGGGGGCGCCCGCCAGCGCGGTTGGGGG 596

RESULT 8
AW028197/c
LOCUS
DEFINITION
IMAGE:2536200 3' similar to TR:060299 O60299 KIAA0552 PROTEIN. ;
mRNA sequence.
AW028197
AW028197.1 GI:5886953
EST.
SOURCE
human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 397)

REFERENCE

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE

Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapsb-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1067 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 139.

FEATURES

Location/Qualifiers

1..397

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2536200"

/clone_lib="Soares_thymus_NHFT"

/dev_stage="fetal"

/lab_host="DH10B (phage-resistant)"

/note="Organ: thymus, pooled; Vector: pT73D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5,

TGTACCAACTGAAGTGGAGCGCGCGCAACGTTTTTTTTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo.

67 a 120 c 124 g 86 t

BASE COUNT

ORIGIN

Query Match 20.5%; Score 366.6; DB 10; Length 397;
Best Local Similarity 95.2%; Pred. No. 4.3e-57;
Matches 378; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 502 TCTGGGCGCTGTGACACTCCGCGCGGAATCTCCATGTCCAGCCTGCCACACAGCACC 561
DB 397 TCTGGGCGCTGTGACAAATCCGCTGGAAATTCATGTCCAGCCTGCCACACAGCACA 338
QY 562 AGCAGCAGCTAATCAGCTGGACCCGCTGTCACACCCGTGGGACCCACAGCGCTTTGGG 621
DB 337 GGCAGCAGCTACCACTGAGTGGACCCGCTGTCACACAGTGGGACCCACAGCGCTTTGGG 278
QY 622 GGCTCCGCGCCACAAATCACCAGCGGATCTCTCTCCAGGACAGCAACATGATGAGCCTG 681
DB 277 GGCTCCGCGCCACAAATCACCAGCGGATCTCTCTCCAGGACAGCAACATGATGAGCTG 218
QY 682 AAGGCTCTGCTCTCTCTCTCGAGCGAGGTAGCAAGCTGGGCCACTCGAACAAGGACAGCAAG 741
DB 217 AAGGCTCTGCTCTCTCTCTCGAGCGAGGTAGCAAGTTGGGCCACTCGAACAAGGACAGCAAG 158
QY 742 GGCCCTCTGCTGTCGCTCCGCTCCCATCTCCACGAGCAGTCCAGCATCCAGAGCTGGAG 801
DB 157 GGCCCTCTGCTGTCGCTCCGCTCCCATCTCCACGAGTAGTGCAGCATCCAGAGCTGGAG 98
QY 802 CAGAAGCTGTTGGAGAGGAGGCGGCCCTCCAGAAGCTGTCAGCGCAGCTTTGAGGAGAG 861
DB 97 CAGAAGCTGTTGGAGAGGAGGCGGCCCTCCAGAAGCTGTCAGCGCAGCTTTGAGGAGAG 38
QY 862 GAGCTTGCTCCAGCTTGCTGCTTACGAGGAGCGGCGC 898
DB 37 GAGCTTGCTCCAGCTTGCTGCTTACGAGGAGCGGCGC 1

RESULT 9

AW016544/c

LOCUS

DEFINITION

UI-H-B10p-abg-g-04-0-UI.s1 NCI_CGAP_Sub2 Homo sapiens cDNA clone

AW016544

344 bp mRNA linear EST 10-SEP-1999

```

ACCESSION      AW016544
VERSION        AW016544.1  GI:5865301
KEYWORDS       EST.
SOURCE         Homo sapiens
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 344)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs@mail.nih.gov
               Oligo-dt track not found, Not I site shown in beginning of sequence
               is likely internal to the message. cDNA Library Preparation: M.B.
               Soares Lab Clone distribution: NCI-CGAP clone distribution
               information can be found through the I.M.A.G.E. Consortium/LLNL at:
               www-bio.llnl.gov/bbrp/image/image.html
               Seq primer: M13 Forward
               POLYA=No.

FEATURES             Location/Qualifiers
     1..344
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:2711982"
     /clone_lib="NCI CGAP Sub2"
     /lab_host="DH10B (Life Technologies)"
     /note="Vector: pT73D-Pac (Pharmacia) with a modified
     polylinker; Site 1: Not I; Site 2: Eco RI; The
     NCI CGAP Sub2 library is a subtracted library derived from
     B1. B1 constitutes a mixture of 21 normalized or
     subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22
     , NCI CGAP Pr28, NCI CGAP Co10, NCI CGAP Co16,
     NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3,
     NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8,
     NCI CGAP CLL1, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5
     , NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6
     , NCI CGAP Brn25. These 21 libraries were pooled and a
     single-stranded DNA preparation of the resulting mixture
     was used as a tracer in a subtractive hybridization with a
     driver whose composition is detailed below: NCI CGAP Kid3
     pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE
     ClonIDs 1322376-1323911, 1456008-1456775, 1500552-1502855
     ) NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725,
     3776-3778 (IMAGE ClonIDs 1322912-1325831,
     1473368-1472903, 1492104-1493255) NCI CGAP Lu5 pool 1 LLAM
     3575-3582, 3851-3854 (IMAGE ClonIDs 1414920-1417991,
     1520904-1522439) NCI CGAP GC4 pool 1 LLAM 3164-3167,
     3716-3720, 3733-3735 (IMAGE ClonIDs 1257096-1258631,
     1469064-1470983, 1475592-1476743) NCI CGAP Pr22 pool 1
     LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE ClonIDs
     985608-986759, 1101192-1101959, 1217928-1220615)
     NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
     ClonIDs 1057416-1061255, 1144584-1145351) The resulting
     subtracted library contained 4 million recombinants.
     Subtraction was performed as previously described [Bonaldo
     , Lennon & Soares (1996): Normalization and Subtraction:
     Two Approaches To Facilitate Gene Discovery. Genome
     Research 6, 791-806.
     TAG LIB=NCI CGAP GC4
     TAG TISSUE=germ Cell
     TAG_SEQ=AAATC"
BASE COUNT      51 a 106 c 112 g 74 t 1 others
ORIGIN
Query Match      18.9%; Score 338.2; DB 10; Length 344;
Best Local Similarity 98.8%; Pred No. 6,3e-52;
Matches 340; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 554 ACAGCAGCAGCAGCTTACCGCTGGACCGCTGGTCCACACCCGTGGGACCCCAAGCC 613
|||||

```

```

Db 344 ACAGCAGCAGCAGCTTACCGCTGGACCGCTGGTCCACACCCGTGGGACCCCAAGCC 285
QY 614 GTTTTGGGGGCTCCGCCACAAATCATCCAGGGCATCGTCTCTCCAGGACAGCAACATGA 673
|||
Db 284 GTTTTGGGGGCTCCGCCACAAATCATCCAGGGCATCGTCTCTNCGAGCAGCAACATGA 225
QY 674 TGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGCCCACTCGAACAAAGG 733
Db 224 TGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGCCCACTCGAACAAAGG 165
QY 734 CAGACAGGGCCCTCGTGTCTCCGCTCCCACTCTCCAGGACAGGTGCGAGCATCCAGG 793
Db 164 CAGACAGGGCCCTCGTGTCTCCGCTCCCACTCTCCAGGATGAGTGCAGCATCCAGG 105
QY 794 AGCTGGACGACAGCTGTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTG 853
Db 104 AGCTGGACGACAGCTGTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTG 45
QY 854 AGGAGAGGAGCTTGCTCTCCAGCTGGCTTACGAGGAGCGGCCG 897
Db 44 AGGAGAGGAGCTTGCTCTCCAGCTGGCTTACGAGGAGCGGCCG 1

RESULT 10
LOCUS    BF549120
DEFINITION UI-R-A0-af-c-04-0-UI.r1 UI-R-A0 Rattus norvegicus cDNA clone
ACCESSION BF549120
VERSION    BF549120.1  GI:11640185
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE  1 (bases 1 to 394)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
           discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    9704477
COMMENT    Contact: Soares, MB
           Program for Rat Gene Discovery and Mapping
           University of Iowa
           451 Eckstein Medical Research Building Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: msoares@blue.weeg.uiowa.edu
           cDNA Library Preparation: M.B. Soares Lab Clone distribution:
           clones will be available through Research Genetics (www.resgen.com)
           This clone is also available through the I.M.A.G.E. Consortium at
           LLNL (info@image.llnl.gov). IMAGE ID= 1791781
           Seq primer: M13 Forward.
           Location/Qualifiers
               1..394
               /organism="Rattus norvegicus"
               /strain="Sprague-Dawley"
               /db_xref="taxon:10116"
               /clone="UI-R-A0-af-c-04-0-UI"
               /clone_lib="UI-R-A0"
               /dev_stage="adult"
               /lab_host="DH10B (Life Technologies)"
               /note="Vector: pT73D-Pac (Pharmacia) with a modified
               polylinker; Site 1: Not I; Site 2: Eco RI; This library
               consists of a mixture of individually tagged normalized
               libraries constructed from rat placenta, adult lung, brain
               , liver, kidney, heart, spleen, ovary, and muscle. The tag
               is a string of 3-5 nucleotides present between the Not I
               site and the oligo-dr track which allows identification of
               the library of origin of a clone within the mixture."
BASE COUNT      89 a 125 c 103 g 77 t
ORIGIN

```



```

source
1. .743
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3638203"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: placenta; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      119 a      229 c      226 g      168 t      1 others
ORIGIN
Query Match      16.7%; Score 298.4; DB 10; Length 743;
Best Local Similarity 97.3%; Pred. No. 1.3e-44;
Matches 324; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy      622 GGCTCCGCCCAACATCACCAGGCGATCGTCTCCAGGACGACCAACATGATGAGCCTG 681
      |||
Db      711 GGGCCGCCGCCAGCAATCACAGG-CATGCTCTCCAGGACGACCAACATGATGAGCCTG 653

Qy      682 AAGGCTGTGCTCTCCGACGGAGGTAGCAAGCTGGGCCACTCGAACAAAGCGACACAAG 741
      |||
Db      652 AA-GCTGTGCTCTCCGACGGAGGTAGCAAGCTGGGCCACTCGAACAAAGCGACACAAG 594

Qy      742 GGGCCCTGTGTGCTCGCTCCCTCCCTCCAGGACGAGTGCGAGATCCAGGAGCTGGAG 801
      |||
Db      593 GGGCCCTGTGTGCTCGCTCCCTCCCTCCAGGACGAGTGCGAGATCCAGGAGCTGGAG 534

Qy      802 CAGAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAGAGTGCAGCGAGCTTTGAGGAGAG 861
      |||
Db      533 CAGAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAGAGTGCAGCGAGCTTTGAGGAGAG 474

Qy      862 GAGCTTGCTCCAGCTCGCTTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAG 921
      |||
Db      473 GAGCTTGCTCCAGCTCGCTTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAG 414

Qy      922 GGGCCGAGGCCAAAGGGCGCAACAGCTCAAG 954
      |||
Db      413 GGGCCGAGGCCAAAGGGCGCAACAGCTCAAG 381

RESULT 13
BE384131/c
LOCUS      601272956F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614076 5',
DEFINITION      mRNA sequence.
ACCESSION      BE384131
VERSION        BE384131.1 GI:9329496
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCFD/PTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at: image.llnl.gov
Plate: LLCN275 row: P column: 13
High quality sequence stop: 670.

FEATURES
source
1. .673
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3614076"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      110 a      211 c      208 g      144 t
ORIGIN
Query Match      16.0%; Score 286.8; DB 10; Length 673;
Best Local Similarity 98.7%; Pred. No. 1.7e-42;
Matches 310; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy      642 CCAGGGATCGTCTCCAGGACGACCAACATGATGAGCTCTGTCTCTCTCCGA 701
      |||
Db      669 CCCAGGATCGTCTCTCCA-GACAGCAACATGATGAGCTCTGTCTCTCTCCGA 611

Qy      702 CGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACAGAGGG-CCCTCTGTGTGCTCGCT 760
      |||
Db      610 CGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACAGAGGGCCCTCTGTGTGCTCGCT 551

Qy      761 CCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTGGAGAGGG 820
      |||
Db      550 CCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTGGAGAGGG 491

Qy      821 AGGGCGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGG 880
      |||
Db      490 AGGGCGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGG 431

Qy      881 CCTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGAGCCCAAGGCG 940
      |||
Db      430 CCTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGAGCCCAAGGCG 371

Qy      941 GCAACAAAGCTCAAG 954
      |||
Db      370 GCAACAAAGCTCAAG 357

RESULT 14
BF706253
LOCUS      280459 MARC 3BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
DEFINITION      BF706253
ACCESSION      BF706253
VERSION        BF706253.1 GI:11997851
KEYWORDS      EST.
SOURCE        cow.
ORGANISM      Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE      1 (bases 1 to 472)
AUTHORS      Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
TITLE        Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL        Genome Res. 11 (4), 626-630 (2001)
MEDLINE
COMMENT      Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

```


Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCAGGCG

Plate: 77 row: A column: 4

Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1. 472

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

92.a 171 c 123 g 86 t

BASE COUNT

ORIGIN

Query Match 16.0%; Score 286; DB 12; Length 472;

Best Local Similarity 92.3%; Pred. No. 2.2e-42;

Matches 301; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ATGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60

Db 147 ATGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 206

Qy 61 TCGCAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCGACGG 120

Db 207 TCGCAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCGACGG 266

Qy 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGTCAGCGCAAGTCAGCTCCCAATGGC 180

Db 267 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGTCAGCGCAAGTCAGCTCCCAATGGC 326

Qy 181 AAGCGCAAGACTTCTTACATCAAGTTCAGCGCAAGAGCGGGCTCCATCACCCA 240

Db 327 AAGCGCAAGACTTCTTACATCAAGTTCAGCGCAAGAGCGGGCTCCATCACCCA 386

Qy 241 GATTACAGGCACTCTCCAGCGGGGATTTAGGGGCCAGGTGGGGTGGACTTTGACCG 300

Db 387 GATTACAGGCACTCTCCAGTGGGACCTAGGGGCCAGGAGTGGACTTTGACCCA 446

Qy 301 TCCACACCCCAAGCTCATGCCCTT 326

Db 447 TCCACCCCAAGCTCATGCCCTT 472

RESULT 15

AZ858490/c

LOCUS

DEFINITION 2M016310R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M016310 R, DNA sequence.

ACCESSION AZ858490

VERSION AZ858490.1 GI:13051689

KEYWORDS GSS.

SOURCE house musculus

ORGANISM

REFERENCE 1 (bases 1 to 658)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL

COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: dduun@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0163 row: L column: 10

Seq primer: CACACAGCAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 658.

FEATURES

source

1. 658

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clones="UUGC2M016310"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 [gi|4732114|gb|AF129072.1], a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 140 a 173 c 177 g 168 t

ORIGIN

Query Match 15.2%; Score 272.4; DB 17; Length 658;

Best Local Similarity 86.7%; Pred. No. 6.9e-40;

Matches 300; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1 ATGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60

Db 582 ATGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 523

Qy 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGG 120

Db 522 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAATCGATCTCAGATGG 463

Qy 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGTCAGCGCAAGTCCAGCTCCAAATGGGC 180

Db 462 CTGCTGAGTTGGCTTCTCCAGAGCTCCCGTCAGCGCGCTGGAAAGTCAAGTTCCAAAATGGGA 403

Qy 181 AAGCGCAAGACTTCTTACATCAAGTTCAGCGCAAGAGCGGGCTCCCATCACCCA 240

Db 402 AAGCGCAAGACTTCTTACATCAAGTTCAGCGCAAGAGCGGGCTCCCATCACCCA 343

Qy 241 GATTACAGGCACTCTCCAGCGGGGATTTAGGGGCCAGGTGGGGTGGACTTTGACCG 300

Db 342 GATTACAGGCACTCTCCAGTGGGACATAGGGGTTCAGCGGAGTAGATTTTCATCCA 283

Qy 301 TCCACACCCCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGG 346

Db 282 GCCACCCCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGG 237

Mon Jun 16 08:15:58 2003

us-09-513-888c-3.rst

Page 11

Search completed: June 15, 2003, 03:41:58
Job time : 2494.59 secs

THIS PAGE BLANK (USPTO)